

## PATENT COOPERATION TREATY

PCT

## **NOTIFICATION OF ELECTION**

(PCT Rule 61.2)

## From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents  
United States Patent and Trademark  
Office  
Box PCT  
Washington, D.C.20231  
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 08 June 2000 (08.06.00)	DROIT DES MARQUES IN INVENTION D'AMERIQUE in its capacity as elected Office
International application No. PCT/SE99/01784	Applicant's or agent's file reference 53371-56822
International filing date (day/month/year) 06 October 1999 (06.10.99)	Priority date (day/month/year) 06 October 1998 (06.10.98)
Applicant <b>ZAPIROPOULOS, Peter, G. et al</b>	

- 1. The designated Office is hereby notified of its election made:**

in the demand filed with the International Preliminary Examining Authority on:

05 May 2000 (05.05.00)

in a notice effecting later election filed with the International Bureau on:

2. The election  was

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer  F. Baechler
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

## PATENT COOPERATION TREATY

PCT

REC'D 24 JAN 2001

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

15

Applicant's or agent's file reference 53371-56822	<b>FOR FURTHER ACTION</b>		See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/SE99/01784	International filing date (day/month/year) 06.10.1999	Priority date (day/month/year) 06.10.1998	
International Patent Classification (IPC) or national classification and IPC7 A61K 48/00, C07K 14/705, A61K 38/17, G01N 33/68			
Applicant Karolinska Innovations AB et al			

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 2 sheets.

3. This report contains indications relating to the following items:

- I  Basis of the report
- II  Priority
- III  Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV  Lack of unity of invention
- V  Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI  Certain documents cited
- VII  Certain defects in the international application
- VIII  Certain observations on the international application

Date of submission of the demand 05.05.2000	Date of completion of this report 12.01.2001
Name and mailing address of the IPEA/SE Patent- och registreringsverket Box 5055 S-102 42 STOCKHOLM Facsimile No. 08-667 72 88	Authorized officer Patrick Andersson/EÖ Telephone No. 08-782 25 00

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

## I. Basis of the report

## 1. With regard to the elements of the international application:\*

 the international application as originally filed the description:

pages 1-23, as originally filed

pages \_\_\_\_\_, filed with the demand

pages \_\_\_\_\_, filed with the letter of \_\_\_\_\_

 the claims:

pages \_\_\_\_\_, as originally filed

pages \_\_\_\_\_, as amended (together with any statement) under article 19

pages \_\_\_\_\_, filed with the demand

pages 1, 2, filed with the letter of 15.12.2000

 the drawings:

pages \_\_\_\_\_, as originally filed

pages \_\_\_\_\_, filed with the demand

pages \_\_\_\_\_, filed with the letter of \_\_\_\_\_

 the sequence listing part of the description:

pages \_\_\_\_\_, as originally filed

pages \_\_\_\_\_, filed with the demand

pages \_\_\_\_\_, filed with the letter of \_\_\_\_\_

## 2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language \_\_\_\_\_ which is:

 the language of a translation furnished for the purposes of international search (under Rule 23.1(b)). the language of publication of the international application (under Rule 48.3(b)). the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

## 3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

 contained in the international application in written form. filed together with the international application in computer readable form. furnished subsequently to this Authority in written form. furnished subsequently to this Authority in computer readable form. The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished. The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.4.  The amendments have resulted in the cancellation of: the description, pages \_\_\_\_\_ the claims, Nos. \_\_\_\_\_ the drawings, sheet/fig \_\_\_\_\_5.  This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2 (c)).\*\*

\* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

\*\* Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

**III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability**

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

- the entire international application,  
 claims Nos. 7 (partially), 8-10 and 16 completely

because:

- the said international application, or the said claims Nos. 7 (partially) and 16 completely  
relate to the following subject matter which does not require an international preliminary examination (*specify*):

Claims 7 partially 16 completely relate to methods of treatment of the human or animal body by therapy/ diagnostic methods practised on the human or animal body See PCT Rule. 67.1.(iv).

- the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 8-10  
are so unclear that no meaningful opinion could be formed (*specify*):

Claim 8 relates to a screening method, it is not clear from this claim or the description how the polypeptide is used to achieve the screening. Claim 9 relates to a method of synthesis a modified drug, it is not clear from the description or the claim how the protein is used. Claim 10 relates to a drug identified by any of the methods in 8 or 9. Thus, claims 8-10 is considered to fail to comply with PCT-article 6 and have not been searched. Moreover, claim 15 concerns a drug per se, the drug may be a known substance.

... / ...

- the claims, or said claims Nos. \_\_\_\_\_ are so inadequately supported  
by the description that no meaningful opinion could be formed.  
 no international search report has been established for said claims Nos. \_\_\_\_\_

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

- the written form has not been furnished or does not comply with the standard.  
 the computer readable form has not been furnished or does not comply with the standard.

**INTERNATIONAL PRELIMINARY EXAMINATION REPORT**

International application No.

PCT/SE99/01784

**Supplemental Box**

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: III

The description or the claims does not give any indication on that any unknown substance(s) that are preferred. For this reason, a search or a statement concerning novelty and an inventive step can not be made for claim 10.

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement****1. Statement**

Novelty (N)	Claims	5-7, 11-15	YES
	Claims	1-4, 17	NO
Inventive step (IS)	Claims		YES
	Claims	1-7, 11-15, 17	NO
Industrial applicability (IA)	Claims	1-7, 11-15, 17	YES
	Claims		NO

**2. Citations and explanations (Rule 70.7)**

The claimed invention relates to nucleic acids, proteins participating in the human PTCH/SHH pathway as well as related vectors, cells, antibodies, methods and uses.

The following document is considered relevant:

D1) Motoyama J et al., "Ptch2, a second mouse Patched gene is co-expressed with sonic hedgehog"

D1 discloses a mouse patched gene and its amino acid sequence having 91% homology in an 1140 amino acid overlap. The expression is analysed with RNA probes. In the present description a comparison between the sequence of D1 and the sequence claimed is made; the differences between the sequences are pointed out and it is speculated what effect these differences may have. However, the difference in function between the two has not been shown. Therefore, the expression "a variant thereof capable of participating in the human PTCH/SHH pathway" is unclear since a person skilled in the art get no information on what is covered e.g. by experiments showing unique functions of the claimed protein. The protein/nucleic acid of D1 would interact(i.e. be capable of participating) in some way also in the human pathway and can therefore be regarded as a variant of SEQ ID NO: 1. Thus, the invention according claims 1-4 and 17 lacks novelty.

....

**INTERNATIONAL PRELIMINARY EXAMINATION REPORT**

International application No.

PCT/SE99/01784

**Supplemental Box**

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: V

Nucleic acids hybridising to a nucleic acid, vectors, recombinant cells, antibodies, kits are all features obvious to a person skilled in the art; whether claims 11-15 involves an inventive step depends on if the claims relate to are inventive.

Thus the invention according to claims 11-15 is novel and industrially applicable, but is not considered to involve an inventive step.

D1 discloses that the patched gene is involved in cancer, that the gene or its product can be used in some way in a medicament or diagnoses is considered obvious to a person skilled in the art. The present application seems not disclose any specific uses in medicament or diagnoses (e.g. a link between a mutation in the patched and a medical state) that goes beyond what a person skilled in the art would expect in view of D1. Thus the invention according to claims 5-7 is novel but not considered to involve an inventive step.

CLAIMS

1. An isolated human protein or an analogue or variant thereof capable of participating in the human PTCH/SHH pathway during embryonic development and/or carcinogenesis comprising at least about 1040 amino acids of SEQ ID NO: 1.
- 5 2. A protein according to claim 1, which is essentially comprised of SEQ ID NO: 1.
3. A nucleic acid encoding a protein according to any one of claims 1 and 2.
4. A nucleic acid encoding a protein according to claim 3 comprising at least about 3094 bases of SEQ ID NO: 2.
- 10 5. An isolated genomic nucleic acid comprising parts or all of SEQ ID NO: 5.
6. An isolated variant of the nucleic acid according to claim 5.
7. A nucleic acid having the sequence of a splicing variant selected from the group of variants defined in Figure 2B and SEQ ID NO:3 and SEQ ID NO:4.
- 15 8. An isolated nucleic acid capable of specifically hybridising to a nucleic acid according to any one of claims 3-7.
9. A protein or polypeptide encoded by a nucleic acid according to claim 7.
10. A protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9, for use as a medicament.
- 20 11. Use of a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 in the manufacture of a medicament for the treatment of a condition involving tumors, such as BSS.
12. A method of in vitro or in vivo diagnosis, wherein a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 is used.
- 25 13. A method of screening wherein a library of suitable candidate compounds is screened for modified drugs using a protein according to claim 1 or 2 or a polypeptide according to claim 9 as a lead compound.
14. A method of synthesis of a modified drug, wherein a protein according to claim 1 or 2, or a protein or polypeptide according to claim 9, is used.
- 30 15. A modified drug identified by the method according to claim 13 or synthesized according to claim 14.

16. A vector comprising a nucleic acid according to any one of claim 3-8.
17. A recombinant cell comprising a vector according to claim 16.
18. An antibody which specifically binds to a protein or polypeptide according to claim 1, 2 or 9.
- 5 19. A recombinant cell expressing an antibody according to claim 18.
20. A kit for the detection of a human PTCH2 gene or polypeptide comprising in a container a molecule selected from the group consisting of a nucleic acid according to any one of claims 3-8, a polypeptide or protein according to claim 1, 2 or 9 or an antibody according to claim 18.
- 10 21. Use of a nucleic acid according to any one of claims 3-7 in gene therapy.
22. Use of nucleic acid according to any one of claims 3-7 as a probe, a primer or a diagnostic reagent.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>A61K 48/00, C07K 14/705, A61K 38/17, G01N 33/68</b>		A1	(11) International Publication Number: <b>WO 00/20037</b>
			(43) International Publication Date: <b>13 April 2000 (13.04.00)</b>
(21) International Application Number: <b>PCT/SE99/01784</b>			(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: <b>6 October 1999 (06.10.99)</b>			
(30) Priority Data: <b>9803393-9M 6 October 1998 (06.10.98) SE</b>			
(71) Applicants (for all designated States except US): KAROLINSKA INNOVATIONS AB [SE/SE]; S-171 77 Stockholm (SE). PHARMACIA & UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): ZAPHIROPOULOS, Peter, G. [GR/SE]; Tullinge Strand 96, S-146 54 Tullinge (SE). UNDÉN, Anne, Birgitte [NO/SE]; Gandviksvägen 3, S-181 62 Djursholm (SE). TOFTGÅRD, Rune [SE/SE]; Sätragårdsvägen 209, S-127 36 Skärholmen (SE). RAH-NAMA, Fahimeh [SE/SE]; Stadshagsvägen 18 3 tr, S-112 50 Stockholm (SE). HOLLINGSWORTH, Robert, E. [US/US]; Glaxo Inc., 5 Moore Drive, Research Triangle Park, NC 27709 (US).			
(74) Agents: BERG, S., A. et al.; Albihns Patentbyrå Stockholm AB, P.O. Box 5581, S-114 85 Stockholm (SE).			

(54) Title: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY

(57) Abstract

The present invention relates to a novel human patched-like gene (PTCH2), which for the first time has been cloned and sequenced. Several alternatively spliced mRNA forms of PTCH2 have been identified, including transcripts lacking segments thought to be involved in sonic hedgehog (SHH) binding and mRNAs with differentially defined 3' terminal exons. Further, the invention also relates to the protein encoded by the present PTCH2 as well as to functional analogues and variants thereof.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY

### Technical field

The present invention relates to novel molecules, such as proteins, polypeptides and nucleotides, involved in the hedgehog signalling pathway with putative involvement in embryonic development and carcinogenesis. The invention also relates to various novel advantageous uses of the molecules according to the invention, e.g. in diagnosis and therapy.

### Background

In the study of the development of cells, fruit flies have extensively been used as a model, as they are less complex than mammalian cells.

Pattern formation takes place through a series of logical steps, reiterated many times during the development of an organism. Viewed from a broader evolutionary perspective, across species, the same sort of reiterative pattern formations are seen. The central dogma of pattern formation has been described (Lawrence and Struhl, 1996). Three interlocking and overlapping steps are defined. Firstly, positional information in the form of morphogen gradients allocate cells into non-overlapping sets, each set founding a compartment. Secondly, each of these compartments acquire a genetic address, as a result of the function of active "selector" genes, that specify cell fate within a compartment and also instruct cells and their descendants how to communicate with cells in neighboring compartments. The third step involves interactions between cells in adjacent compartments, initiating new morphogen gradients, which directly organize the pattern.

Taking these steps in greater detail, one finds the first step in patterning to be the definition of sets of cells in each primordium. Cells are allocated according to their positions with respect to both dorsoventral and anterior/posterior axes by morphogen gradients. Allocation of cells in the dorsoventral axis constitutes the germ layers, such as mesoderm or neurectoderm.

In segmentation, the second step (the specification of cell fate in each compartment) is carried out by the gene *engrailed* and elements of the bithorax complex. *Engrailed* defines anterior and posterior compartments both in segmentation and in limb specification.

5

The third step in pattern formation, secretion of morphogens, functions to differentiate patterns within compartments (and thereby establish segment polarity). Initially, all cells within a compartment are equipotent, but they become diversified to form pattern. Pattern formation depends on gradients of morphogens, gradients initiated along compartment boundaries. Such gradients are established by a short-range signal induced in all the cells of the compartment in which the above mentioned selector gene *engrailed* is active. For segment polarity, this signal is Hedgehog. In the adjacent compartment the selector gene is inactive, ensuring that the cells are sensitive to the signal. The Hedgehog signal range is probably only a few rows of cells wide; responding cells become a linear source of a long-range morphogen, that diffuses outward in all directions. There are three known Hedgehogs, Sonic (SHH), Indian (IHH) and Desert (DHH). The proteins they encode can substitute each for each other, but in wildtype animals, their distinct distributions result in unique activities. SHH controls the polarity of limb growth, directs the development of neurons in the ventral neural tube and patterns somites. IHH controls endochondral bone development and DHH is necessary for spermiogenesis. Vertebrate hedgehog genes are expressed in many other tissues, including the peripheral nervous system, brain, lung, liver, kidney, tooth primordia, genitalia and hindgut and foregut endoderm.

25 Thus, segment polarity genes have been identified in flies as mutations, which change the pattern of structures of the body segments. Mutations in these genes cause animals to develop the changed patterns on the surfaces of body segments, the changes affecting the pattern along the head to tail axis. For example, mutations in the gene *patched* cause each body segment to develop without the normal structures in the center of each segment. Instead there is a mirror image of the pattern normally found in the anterior segment. Thus, cells in the center of the segment make the

wrong structures, and point them in the wrong direction with reference to the overall head-to-tail polarity of the animal.

About sixteen genes in the class are known. The encoded proteins include kinases, transcription factors, a cell junction protein, two secreted proteins called wingless (WG) and the above mentioned Hedgehog (HH), a single transmembrane protein called patched (PTC) and some novel proteins not related to any known protein. All of these proteins are believed to work together in signaling pathways that inform cells about their neighbors in order to set cell fates and polarities.

10

PTC has been proposed as a receptor for HH protein based on genetic experiments in flies. A model for the relationship is that PTC acts through a largely unknown pathway to inactivate both its own transcription and the transcription of the *wingless* segment polarity gene. This model proposes that HH protein, secreted from adjacent cells, binds to the PTC receptor, inactivates it and thereby prevents PTC from turning off its own transcription or that of *wingless*. A number of experiments have shown coordinate events between PTC and HH.

Human *patched* gene (PTCH) was recently identified as the gene responsible for the nevoid basal cell carcinoma syndrome (NBCCS), also known as the Gorlin Syndrome, which is an autosomal dominant disorder that predisposes to both cancer and developmental defects (Gorlin (1995) *Dermatologic Clinics* 13:113-125) characterized by multiple basal cell carcinomas (BCCs), medulloblastomas and ovarian fibromas as well as numerous developmental anomalies (Hahn, H., Wicking, C., Zaphiropoulos, P.G., Gailani, M.R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Undén, A.B., Gillies, S., Negus, K., Smyth, I., Pressman, C., Lefell, D.J., Gerrard, B., Goldstein, A.M., Dean, M., Toftgård, R., Chenevix-Trench, G., Wainright, B. and Bale, A.E. (1996): "Mutations of the human homolog of *Drosophila patched* in the nevoid basal cell carcinoma syndrome", *Cell* 85, 841-851; and Johnson, R.L., Rothman, A.L., Xie, J., Goodrich, L.V., Bare, J.W., Bonifas, J.M., Quinn, A.G., Myers, R.M., Cox, D.R., Epstein, E.H. Jr and Scott, M.P.

(1996): "Human homolog of patched, a candidate gene for the basal cell nevus syndrome", *Science* 272, 1668-1671). PTCH codes for a membrane receptor of the autolytically cleaved (protein spliced), amino terminal domain of *sonic hedgehog* (SHH) (Mariago, V., Davey, R.A., Zuo, Y., Cunningham, J.M. and Tabin, C.J. 5 (1996): "Biochemical evidence that patched is the Hedgehog receptor", *Nature* 384, 176-179; and Stone, D.M., Hynes, M., Armanini, M., Swanson, T.A., Gu, Q., Johnson, R.L., Scott, M.P., Pennica, D., Goddard, A., Phillips, H., Noll, M., Hooper, J.E., de Sauvage, F. and Rosenthal, A. (1996): "The tumor-suppressor gene patched encodes a candidate receptor for Sonic hedgehog", *Nature* 384, 129-134). In the 10 non-signalling state, PTCH is thought to inhibit the consecutive signalling of another membrane protein, *smoothened* (SMO), however binding of SHH to PTCH relieves this inhibition (Goodrich, L.V., Milenkovic, L., Higgins, K.M. and Scott, M.P. (1997): "Altered neural cell fates and medullablastom in mouse patched mutants", *Science* 277, 1109-1113). This cascade of signalling events, best characterized 15 in *Drosophila*, also involves a number of intracellular components including *fused* (a serine threonine kinase), *suppressor of fused*, *costal 2*, and *cubitus interruptus* (Ruiz i Altaba, A.: "Catching a Gli-mpse of Hedgehog" (1997) *Cell* 90, 193-196). The latter is a transcription factor that positively regulates the expression of target genes which also include PTCH itself.

20

Mutations in the PTCH gene have been identified in both sporadic and familial BCCs (Gailani, M.R., Stähle-Bäckdahl, M., Leffell, D.J., Glynn, M., Zaphiropoulos, P.G., Pressman, C., Undén, A.B., Dean, M., Brash, D. E., Bale, A.E. and Toftgård, R. (1996): "The role of human homologue of *Drosophila* patched in sporadic 25 basal cell carcinomas" *Nature Genet.* 14, 78-81). The lack of the normal PTCH protein in these cells allows the constitutive signalling of SMO to occur, resulting in the accumulation of mutant PTCH mRNAs (Undén, B. A., Zaphiropoulos, P.G., Bruce, K., Toftgård, R., and Stähle-Bäckdahl, M. (1997): "Human patched (PTCH) mRNA is overexpressed consistently in tumor cells of both familial and sporadic 30 basal cell carcinoma", *Cancer Res.* 57, 2336-2340).

WO 96/11260 discloses the isolation of *patched* genes and the use of the PTC protein to identify ligands, other than the established ligand Hedgehog, that bind thereto.

- 5 However, there is still a need of a further understanding of the SHH/PTCH cell signalling, which may be provided by disclosure of further genes, peptides and proteins involved therein.

Summary of the invention

10 The present invention provides a significant step forward regarding the understanding of the above described pathway. By a combination of cDNA library and RACE analysis a novel human *patched*-like gene (PTCH2) has been cloned and sequenced. Several alternatively spliced mRNA forms of PTCH2 have been identified, including transcripts lacking segments thought to be involved in sonic hedgehog (SHH) binding and mRNAs with differentially defined 3' terminal exons. Accordingly, the invention relates to isolated such mRNAs as well as to cDNAs complementary thereto.

Brief description of the drawings

20 Figure 1 shows the genomic sequence of SEQ ID NO:5, wherein exons and introns are designated in the genomic sequence of the novel human patched 2 gene.

Figure 2A discloses an amino acid sequence comparison of the human PTCH2 (upper lines) and PTCH1 (lower lines) sequences.

25 Figure 2B is a representation of the alternative splicing events that result in different C-termini.

Figure 2C is a representation of the different variations of spliced transcripts encompassing exon 1 and exon 2 sequences.

30 Figure 3A is a dark-field photomicrograph of a BCC tumor hybridised with <sup>35</sup>S-labeled antisense probe showing abundant signal for PTCH1 mRNA (light grains) in all BCC tumor cells.

Figure 3B discloses PTCH2 mRNA overexpression in BCC and is in contrast mainly expressed in the basaloid cells in the periphery of the tumor nests.

Figure 3C is another BCC showing a strong PTCH2 mRNA signal in the periphery of the tumor nest (Tu), whereas no signal is detected in epidermis (Ep).

5 Figure 3D are sections of the same tumor (C) hybridised with the PTCH2 sense probe showed no signal.

Figure 3E shows immunoreactivity for Ki-67.

Figure 3F discloses how tumor nests under high power magnification demonstrate abundant PTCH2 mRNA signal (black grains) in the dark basaloid tumor cells and  
10 lower signal in the center (arrow).

#### Definitions

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

15 The terms "isolated", "purified" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany it as found in its native state.

20 The term "nucleic acid" refers to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, encompasses known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides.

25 A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include  $^{32}\text{P}$ , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in a ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or mono-

clonal antibodies are available (e.g., the peptide of SEQ ID NO:1 can be made detectable, e.g., by incorporating a radio-label into the peptide, and used to detect antibodies specifically reactive with the peptide).

5 As used herein a "nucleic acid probe" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.* A, G, C, or T) or modified bases (7-deazaguanosine, inosine, *etc.*) In addition, the bases in a  
10 probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridisation. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes  
15 may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridisation conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphore, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

20 A "labeled nucleic acid probe" is a nucleic acid probe that is bound, either covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

25 The term "target nucleic acid" refers to a nucleic acid (often derived from a biological sample), to which a nucleic acid probe is designed to specifically hybridise. It is either the presence or absence of the target nucleic acid that is to be detected, or the amount of the target nucleic acid that is to be quantified. The target nucleic acid has  
30 a sequence that is complementary to the nucleic acid sequence of the corresponding probe directed to the target. The term target nucleic acid may refer to the specific

subsequence of a larger nucleic acid to which the probe is directed or to the overall sequence (e.g., gene or mRNA) whose expression level it is desired to detect. The difference in usage will be apparent from context.

- 5      The term "recombinant" when used with reference to a cell, or nucleic acid, or vector, indicates that the cell, or nucleic acid, or vector, has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid, or that the cell is derived from a cell so modified.
- 10     The term "identical" in the context of two nucleic acids or polypeptide sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2: 482, by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, GESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI) or by inspection. The BLAST algorithm performs a statistical analysis of the similarity between two sequences; see e.g., Karlin and Altschul (1993) *Proc. Nat'l Acad. Sci. USA* 90: 5873-5787.

- 15     The term "substantial identity" or "substantial similarity" in the context of a polypeptide indicates that a polypeptides comprises a sequence with at least 70% sequence identity to a reference sequence, or preferably 80%, or more preferably 85% sequence identity to the reference sequence, or most preferably 90% identity over a comparison window of about 10-20 amino acid residues. An indication that two polypeptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a polypeptide 20    is substantially identical to a second polypeptide, for example, where the two peptides differ only by a conservative substitution.

5

An indication that two nucleic acid sequences are substantially identical is that the polypeptide which the first nucleic acid encodes is immunologically cross reactive with the polypeptide encoded by the second nucleic acid. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridise to each other under stringent conditions.

10

The phrase "hybridising specifically to", refers to the binding, duplexing, or hybridising of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular DNA or RNA. The term "stringent conditions" refers to conditions under which a probe will hybridise to its target subsequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridise specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point <sup>TM</sup> for the specific sequence at a defined ionic strength and pH. The T<sub>m</sub> is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridise to the target sequence at equilibrium. (As the target sequences are generally present in excess, at T<sub>m</sub>, 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

20

25

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, *etc.*; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

The term "immunoassay" is an assay that utilizes an antibody to specifically bind an analyte. The immunoassay is characterized by the use of specific binding properties of a particular antibody to isolate, target, and/or quantify the analyte.

The phrases "specifically binds to a protein" or "specifically immunoreactive with", when referring to an antibody refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind preferentially to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbour Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

A "gene product", as used herein, refers to a nucleic acid whose presence, absence, quantity, or nucleic acid sequence is indicative of a presence, absence, quantity, or nucleic acid composition of the gene. Gene products thus include, but are not limited to, and mRNA transcript acDNA reverse transcribed from an mRNA, and RNA

transcribed from that cDNA, a DNA amplified from the cDNA, an RNA transcribed from the amplified DNA or subsequences of any of these nucleic acids. Polypeptides expressed by the gene or subsequences thereof are also gene products. The particular type of gene product will be evident from the context of the usage of the  
5 term.

A "modified drug" means a compound, which retains the pharmaceutical properties of the original drug or active substance while the structure thereof has been modified. Further, encompassed by the term "drug" are also compounds useful in diagnostic methods by their specific binding properties.  
10

#### Detailed description of the invention

In a first aspect, the present invention relates to an isolated human protein, or an analogue or a variant thereof, capable of participating in the human PTCH/SHH pathway during embryonic development and/or carcinogenesis, such as basal cell carcinoma. The novel protein according to the invention is encoded by a novel gene, which isolated nucleic acid is described in detail below and which is denoted *patched 2* (PTCH2) due to its similarities with *patched 1* (PTCH1). Accordingly, the protein according to the invention exhibits substantial differences in sequence and functions when compared to human PTCH1protein. The protein according to the invention is best characterized by its functions which when compared to human PTCH1 are similar but distinct therefrom in certain ways, more specifically disclosed below in the section "Results and discussion". The novel human PTCH2 protein according to the invention is also distinct from the previously isolated mouse PTCH2. Thus, in the preferred embodiment thereof, it comprises a substantial part of the amino acid sequence disclosed in SEQ ID NO: 1 and submitted to the Gen-Bank under protein id no AAD17260.1. even though it is to be understood that the present invention encompasses any fragment, analogue or variant thereof exhibiting the biological functions of the PTCH2 protein disclosed herein. Thus, preferably, the present protein comprises at least about 1000, more preferably at least about  
15  
20  
25  
30

1040 and most preferably essentially all of the amino acids of the sequence denoted SEQ ID NO: 1, such as about 1100.

5       The proteins according to the invention are easily prepared by someone skilled in this field by recombinant DNA techniques using the molecules disclosed below or any synthetic method (*see e.g.* Barany and Merrifield, Solid-Phase Peptide synthesis, pp. 3-284 in *The Peptides: Analysis, Synthesis, Biology, Vol. 2: Special Methods in Peptide synthesis, Part A*, Merrifield *et al.*, J. Am. Chem. Soc., 2149-2156).

10      The present invention also relates to the use of the peptides, polypeptides and proteins disclosed herein as lead compounds in methods aimed at finding novel substances, i.e. modified drugs, such as substances exhibiting equivalent or even more advantageous properties than the lead compounds as such. Such modified drugs may 15     also be designed by methods of combinatorial chemistry, wherein a structurally similar compound is specifically designed *e.g.* by aid of computers. Alternatively, the present modified drug is identified by screening of a library of candidate compounds, *e.g.* using an antibody according to the invention. In the present context, it is to be understood that when such a modified drug has been identified, it is possible 20     to produce it by any other suitable technique. The invention also relates to proteomic methods wherein the present molecules are used as well as to such a use *per se*.

25      A second aspect of the present invention is a nucleic acid encoding a protein, an analogue or a variant thereof as defined above, that is, the protein coding region of the novel human isolated PTCH2 gene. The PTCH2 gene is 57% identical to PTCH1 and 91% identical to the published mouse Ptch2 sequence (*see Motoyama et al., (1998), supra*). Thus, preferably, the nucleic acid according to the present invention comprises at least about 3000 bases, more preferably at least about 3094 bases and most preferably essentially all of the sequence denoted SEQ ID NO: 2.

In a specific aspect, the present invention relates to the isolated human genomic PTCH2 nucleic acid comprising parts or all of the genomic sequence denoted SEQ ID NO: 5. In the disclosure of the genomic sequence shown in Fig 1, the exon/intron structure of the present gene is shown. Further to the exons shown therein, exon 12a and 12b has also been identified, as specifically defined by SEQ ID NO:3 and SEQ ID NO:4, respectively. Interestingly, there is a splice variant that joins exon 12a to a 3' segment of exon 12b with conservation of the intronic GT-AG dinucleotides. Exons 12a and 12b are not variants, but the actual exons of the gene identified by sequencing the corresponding genomic region. (Materials and methods were as described below). Accordingly, these findings show that PTCH2 has the same intron/exon structure organization as PTCH1. In another embodiment of this aspect, the present invention relates to a transcript that has skipped only one of the exons 9 and 10 defined in Fig 1. In an alternative embodiment, the transcript according to the invention has skipped both of exon 9 and 10. The splice variants of the present gene are discussed in more detail below in the section "Results", all of which are included within the scope of the present invention. This aspect of the invention advantageously enables design of suitable PCR primers, which in turn enables screening for mutations of all of the coding sections thereof, e.g. by SSCP analysis, sequencing, or any other suitable method known to someone skilled in this field. Thus, the novel human PTCH2 gene according to the invention has been localized by radiation hybrid mapping to chromosome 1p32-35 with D1S211 and WI-1404 as closest flanking markers and with an estimated localization 5.5cR from D1S443. This region is often lost by LOH in various different tumor types, such as neuroblastoma, melanoma, breast cancer, colon cancer etc. Accordingly, PTCH2 is a candidate for a tumor suppressor gene in this region and the present invention also encompass diagnostic methods based on this new disclosure.

To this chromosomal region, three cancer predisposition syndromes have also been mapped, namely, familial melanoma CMM1, modifier locus for familial adenomatous polyposis hMOM1 and Michelin Tire Baby Syndrome. PTCH2 is further a candidate for the gene behind these hereditary syndromes. The present molecules are the-

before advantageously used in the context of these conditions, e.g. in therapy and/or diagnosis, such as in assays.

Further, the invention also relates to various PCR primers based on intronic sequences, allowing amplification of all coding sequence. Such primers are advantageously used for mutation screening.  
5

Further, the present invention also relates to the any isolated nucleic acid capable of specifically hybridising to a nucleic acid according to the invention. In addition, the invention also relates to such an isolated nucleic acid which comprises one or more mutations compared to the genomic sequence as well as the use of the novel isolated nucleic acids, e.g. to identify mutations for diagnostic and/or therapeutic purposes.  
10

Further embodiments of this aspect of the invention includes nucleic acid probes, e.g. DNA probes, labelled nucleic acids, cDNAs, RNAs etc., that is, all gene products obtainable by someone skilled in this field based on the novel isolated human PTCH2 gene.  
15

Another aspect of the invention is a nucleic acid corresponding to any one of the splicing variants disclosed in Figure 2B, a protein or polypeptide encoded thereof as well as various uses thereof.  
20

As regards the preparation of nucleic acids according to the invention, any suitable recombinant DNA technique or synthetic method may be used. (For general laboratory procedures useful in this context, see e.g. Sambrook *et al.*, Molecular Cloning, A Laboratory Manual, 2<sup>nd</sup> ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol. 153, Academic Press, Inc., San Diego, CA; Current Protocols in Molecular Biology, F.M. Ausbel *et al.*, eds., Current Protocols  
25 (1994)).  
30

A further aspect of the present invention is a vector comprising a nucleic acid as defined above. Vectors are e.g. useful for transforming cells in vitro or in vivo to express the proteins and peptides according to the invention and may e.g be plasmids, viruses etc.

5

Another aspect of the invention is a recombinant cell, such as a eucaryotic, e.g. a mammalian cell, or a procaryotic cell, e.g. a bacteria, comprising a vector as defined above. Such cells may e.g. be used to monitor expression levels of the proteins and polypeptides according to the invention in a wide variety of contexts. For example, 10 when the effects of a drug is to be determined, the drug will be administered to the transformed organism, tissue or cell. Accordingly, model systems including such cells are another aspect of the invention.

10

A further aspect of the invention is an antibody, such as a monoclonal or polyclonal 15 antibody, which specifically binds to a protein or polypeptide according to the invention. An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more 20 amino acids primarily responsible for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains, respectively.

20

The invention also encompasses chimeric or other antibodies that binds the present 25 proteins or polypeptides. Further, the invention also relates to the use of the present antibodies in assays. (In this context, see e.g. *Fundamental Immunology*, Third Edition, W.E. Paul, ed., Raven Press, N.Y. 1993).

30

Further, the invention also relates to a recombinant cell expressing an antibody according to the invention.

In general, prokaryotes can be used for cloning the DNA sequences encoding a human anti-PTCH2 immunoglobulin chain. *E. coli* is one prokaryotic host particularly useful for cloning the DNA sequences of the present invention. Microbes, such as yeast are also useful for expression. *Saccharomyces* is a preferred yeast host, with suitable vectors having expression control sequences, an origin of replication, termination sequences and the like as desired. Typical promoters include 3-phosphoglycerate kinase and other glycolytic enzymes. Inducible yeast promoters include, among others, promoters from alcohol dehydrogenase 2, isocytochrome C, and enzymes responsible for maltose and galactose utilization.

10

Mammalian cells are a particularly preferred host for expressing nucleotide segments encoding immunoglobulins or fragments thereof (see, e.g. Winnacker, *From Genes to Clones*, VCH Publishers, N.Y., 1987). A number of suitable host cell lines capable of secreting intact heterologous proteins have been developed in the art, and include CHO cell lines, various COS cell lines, HeLa cells, L cells and myeloma cell lines. Preferably, the cells are nonhuman. Expression vectors for these cells can include expression control sequences, such as an origin of replication , a promoter, an enhancer (Queen *et al.* (1986) *Immunol. Rev.* 89:49), and necessary processing information sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites, and transcriptional terminator sequences. Preferred expression control sequences are promoters derived from endogenous genes, cytomegalovirus, SV40, adenovirus, bovine papillomavirus, and the like (see, e.g., Co *et al.* (1992) *J. Immunol.* 1458: 1149).

25

An additional aspect of the present invention is a kit for the detection of a human PTCH2 gene or polypeptide comprising in a container a molecule selected from the group consisting of a nucleic acid, a polypeptide or a protein or an antibody according to the invention. Further suitable components of such a kit are easily determined by someone skilled in this field as are the conditions for the use thereof.

30

Further, the invention also relates to the use of a nucleic acid selected from the group consisting of SEQ ID NOS: 2-4 and SEQ ID NO: 5 in gene therapy. In addition to said specifically disclosed sequences, any one of the herein disclosed exons may be used to this end. For a review of gene therapy procedures, see Anderson, 5 *Science* (1992) 256:808-813; Nabel and Felgner (1993) *TIBTECH* 11: 211-217; Mitani and Caskey (1993) *TIBTECH* 11: 162-166; Mulligan (1993) *Science* 926-932; Dillon (1993) *TIBTECH* 11: 167-175; Miller (1992) *Nature* 357: 455-460; Van Brunt (1988) *Biotechnology* 6(10): 1149-1154; Vigne (1995) *Restorative Neurology and Neuroscience* 8: 35-36; Kremer and Perricaudet (1995) *British Medical Bulletin* 10 51(1) 31-44; Haddada *et al.* (1995) in *Current Topics in Microbiology and Immunology* Doerfler and Böhm (eds) Springer-Verlag, Heidelberg Germany; and Yu *et al.*, *Gene Therapy*(1994) 1:13-26.

Delivery of the gene or genetic material into the cell is the first critical step in gene 15 therapy treatment of disease. A large number of delivery methods are well known to those of skill in the art. Such methods include, for example liposome-based gene delivery (Debs and Zhu (1993) WO 93/24640; Mannino and Gould-Fogerite (1988) *BioTechniques* 6(7): 682-691; Rose U.S. Pat No. 5,279,833; Brigham (1991) WO 91/06309; and Felgner *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 7413-7414), and 20 replication-defective retroviral vectors harboring a therapeutic polynucleotide sequence as part of the retroviral genome (see, e.g., Miller *et al.* (1990) *Mol. Cell. Biol.* 10:4239 (1990; Kolberg (1992) *J. NIH Res.* 4:43, and Cornetta *et al.* *Hum. Gene Ther.* 2:215 (1991)). Widely used retroviral vectors include those based upon murine 25 leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations thereof. See, e.g., Buchscher *et al.* (1992) *J. Virol.* 66 (5) 2731-2739; Johann *et al.* (1992) *J. Virol.* 66 (5):1635-1640 (1992); Sommerfelt *et al.*, (1990) *Virol.* 176:58-59; Wilson *et al.* (1989) *J. Virol.* 63:2374-2378; Miller *et al.*, *J Virol.* 65:2220-2224 (1991); Wong-Staal *et al.*, PCT/US94/05700, and Rosenberg and Fauci (1993) 30 in *Fundamental Immunology, Third Edition* Paul (ed) Raven Press, Ltd., New York and the references therein, and Yu *et al.*, *Gene Therapy* (1994) supra.

The present invention may also be used in the pharmaceutical industry. For example, it will provide information that eventually may enable cells from fetal tissue, which may be transplanted into patients suffering from e.g. Parkinson's disease or cancer, such as BCC. (For a brief review of methods of drug delivery, see Langer 5 249:1 527-1533 (1990), Remington's Pharmaceutical Sciences, Mack Publishing Company, Philadelphia, PA, 17<sup>th</sup> ed. (1985) etc.)

Detailed description of the drawings

Figure 1 shows the genomic sequence of SEQ ID NO:5, wherein exons and introns 10 are designated in the genomic sequence of the present human patched 2 gene. However, exons 12a and 12b discussed above are not specifically shown in Figure 1, but instead disclosed as the separate sequences SEQ ID NO:3 and SEQ ID NO:4, respectively. Figure 2A discloses an amino acid sequence comparison of the human PTCH2 (upper lines) and PTCH1 (lower lines) sequences. Vertical lines indicate 15 identical amino acids, while dots similar amino acids. The PTCH2 sequence presented is composed of the original cDNA clones and of the products of the 5' RACE analysis.

Figure 2B is a representation of the alternative splicing events that result in different C-termini. In the parotid gland and the colon, the penultimate and the last exon 20 are canonically joined together. In fetal brain however the penultimate exon with part of the 3' intron functions as the terminal exon. The intronic sequence is shown by small letters with the flanking exonic by capital letters. Above the nucleotide sequence, the deduced amino acid sequence is shown, and below is the corresponding 25 sequence of the mouse Ptch2. The conserved intronic dinucleotides are shown by bold letters and the termination signals are indicated by asterisks. Note the absence of conservation of the position of the termination codons between the mouse and human PTCH2 sequences. The putative polyadenylation signals are also shown in this diagram. The genomic organization was obtained by analyzing BAC clones encompassing the PTCH2 gene.

Figure 2C is a representation of the different variations of spliced transcripts encompassing exon 1 and exon 2 sequences. The canonical exons 1 and 2 are shown by boxes and the intron between them by a solid line. The GT and AG dinucleotides spanning the sequences that are used as introns in individual transcripts are indicated by small letters. G, Genomic structure, derived from sequencing segments of BAC clones encompassing the PTCH2 gene; C, Canonical transcript; A, Transcript A (the skipped exons 9 and 10 of this product are not shown in the diagram); B, Transcript B.

Figure 3A is a dark-field photomicrograph of a BCC tumor hybridised with <sup>35</sup>S-labeled antisense probe showing abundant signal for PTCH1 mRNA (light grains) in all BCC tumor cells.

Figure 3B discloses PTCH2 mRNA overexpression in BCC and is in contrast mainly expressed in the basaloid cells in the periphery of the tumor nests.

Figure 3C is another BCC showing a strong PTCH2 mRNA signal in the periphery of the tumor nest (Tu), whereas no signal is detected in epidermis (Ep).

Figure 3D are sections of the same tumor (C) hybridised with the PTCH2 sense probe showed no signal.

Figure 3E shows immunoreactivity for Ki-67 (brown precipitate) seen in the periphery, in the cells that showed strong upregulation of PTCH2 mRNA.

Figure 3F discloses tumor nests under high power magnification demonstrate abundant PATCH2 mRNA signal (black grains) in the dark basaloid tumor cells and lower signal in the center (arrow). Bars (A-E), 24 µm, and F, 6 µm.

EXPERIMENTALMaterials and methods

In the present context, a general reference is made to G. Zaphiropoulos et al., Cancer Res., vol. 59, p. 787-792, February 15, 1999, disclosing useful methods in the present context. All references mentioned in the present application are hereby included herein by reference. The examples below are not intended to limit the scope 5 of the invention but merely as an illustration.

- The RACE analysis was performed essentially as described before (Zaphiropoulos, 10 P.G. and Toftgård, R. (1996): "cDNA cloning of a novel WD repeat protein mapping to the 9q22.3 chromosomal region", DNA Cell Biol. 15, 1049-1056) using the Marathon kit (Promega). The primer sequences used for RACE are available upon request.
- 15 The PTCH2, 35S-labeled RNA probes used for the *in situ* hybridisations, that were performed as previously described (Undén *et al.*, (1997), supra), corresponded to positions 218 to 437 and 838 to 920 in the PTCH2 sequence of SEQ ID NO:1 .

Results and discussion

20 In order to identify additional components of the PTCH/SHH cascade of signalling events, the Incyte LifeSeq™ database (Incyte Pharmaceuticals Inc., Palo Alto, CA, USA) was searched using PTCH sequences. In addition to clones representing the PTCH cDNA, two nearly identical cDNAs were identified, from the parotid gland and the colon, that contained sequences similar to, but distinct from, the 3' end of 25 PTCH. By 5' RACE analysis using fetal brain cDNAs additional sequence information from these transcripts (termed PTCH2) and corresponding to a full length cDNA, was obtained (Fig. 2A ). PTCH2 is 57% identical to PTCH1, with a significantly variable region present between the transmembrane domains 6 and 7, and 30 91% identical to the recently published mouse Ptch2 sequence (Motoyama, J., Takabatake, T., Takeshima, K. and Hui, C. (1998): "Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog", Nature Genet. 18, 104-106). In simila-

5              rity with the mouse gene, PTCH2 lacks the C-terminal extension present in human, mouse and chicken PTCH1 (Goodrich, L.V., Johnson, R.L., Milenkovic, L., McMahon, J.A., and Scott, M.P. (1996): "Conservation of the hedgehog/patched signalling pathway from flies to mice: Induction of a mouse patched gene by Hedgehog", Genes Dev. 10, 301-312, Marigo, V., Scott, M.P., Johnson, R.L., Goodrich, L.V. and Tabin, C.J. (1996): "Conservation in hedgehog signalling: Induction of a chicken patched homolog by Sonic hedgehog in the developing limb", Development 122, 1225-1233). However, according to the present invention, it has been shown  
10              that the human PTCH2 cDNA terminates 36 amino acids earlier than the mouse Ptch2 sequence. Moreover, when 3' RACE was performed from fetal brain, an alternate C-terminal region was identified. This had a high structural similarity with the mouse Ptch2 C-terminal sequence and originates from the genomic region that links the last two exons of PTCH2 (Fig. 2B). Therefore, in these alternatively spliced transcripts, the penultimate exon with a segment of the contiguous 3' intron serves  
15              as the terminal exon.

20              Moreover the human and mouse transcripts differed in the position of the termination signals (the human sequence is 21 amino acids longer), suggesting a non-conserved, species-specific function of this alternate C-terminal domain. The finding of two possible C-terminal regions for PTCH2 is intriguing and implies a role  
25              of this phenomenon in modulating signalling. Additional alternatively spliced transcripts were also identified by the RACE analysis (Fig. 2C). Transcript A lacks the sequence that corresponds to exons 9 and 10 of PTCH1 (preliminary comparisons of the intron-exon junctions of PTCH2 with PTCH1 indicate a similar genomic organization), with the open reading frame being retained at the exon 8 to exon 11 junction. Exons 9 and 10 code for the last part of the first extracellular loop and for transmembrane domains 2 and 3 in the putative structure of the PTCH1 protein.  
30              Furthermore this transcript also lacks a 5' segment of the canonical exon 2, due to the use of an alternative 3' splice site present in this exon, with the open reading frame being maintained. The functional consequence of this alternative splicing is not yet known, but it is interesting to note that the extracellular loops in PTCH1 are

presumed to be involved in binding of the ligand SHH (Marigo *et al.*, (1996), *Nature* **384**, *supra*; Stone *et al.*, (1996), *Nature* **384**, *supra*) and that insertion of a neocassette in intron 9 of the mouse PTCH1 gene is associated with a severe phenotype (Hahn, H., Wojnowski, L., Zimmer, A.M., Hall, J., Miller, G. and Zimmer, A. 5 (1998): "Rhabdomyosarcomas and radiation hypersensitivity in a mouse model of Gorlin syndrome", *Nature Med.* **4**, 619-622). Furthermore, exons 9 and 10 encode part of a putative sterol sensing domain (Osborne, T.F. and Rosenfeld, J.M. (1998): "Related membrane domains in proteins of sterol sensing and cell signalling provide 10 a glimpse of treasures still buried within the dynamic realm of intracellular metabolic regulation", *Curr. Opin. Lipidol.* **9**, 137-140, also found in PTCH1, and which has recently been implicated in mediating the potent modulating effect of cholesterol on SHH/PTCH signalling (Cooper, M.K., Porter, J.A., Young, K.E., and Beachy, P.A. (1998): "Teratogen-mediated inhibition of target tissue response to Shh 15 signalling", *Science* **280**, 1603-1607). Thus, if PTCH2 also serves as a receptor for SHH and/or related factors, the receptor form lacking exons 9 and 10 may show altered signalling properties. Transcript B contains additional sequences between canonical exons 1 and 2, that originate from the 5' end of intron 1. The open reading frame that includes the initiator methionine of exon 1 is not maintained in this transcript, suggesting that, if this transcript is functional, either the methionine in 20 exon 2 or non-methionine codons are used in order to produce a protein product, in similarity to what has been proposed for the alternative spliced products of human PTCH1 (Hahn *et al.*, *Cell* **85**, *supra*). By radiation hybrid mapping the PTCH2 gene was localized to the short arm of chromosome 1, in difference to PTCH1 residing 25 on chromosome 9q22.3.

25

The mouse and zebrafish homologs of PTCH2 have been reported to be expressed in a partly overlapping pattern with PTCH1 during embryonic development and to be induced by SHH (Motoyama *et al.*, (1998) *Nature Genet.* **18**, *supra*, Concorde, J.P., Lewis, K.E., Moore, J.W., Goodrich, L.V., Johnson, R.L., Scott, M.P., and Ingham, P.W. (1996): "Spatial regulation of a zebrafish patched homologue reflects 30

the roles of sonic hedgehog and protein kinase A in a neural tube and somite patterning”, Development 122, 2835-2846), implicating a role in this signalling pathway. We were with this background interested to analyze the expression of PTCH2 in BCCs which show consistent upregulation of PTCH1 in all tumor cells (Undén *et al.*, (1997) Cancer res. 57, supra). In situ hybridisation was performed on six familial and four sporadic BCCs of different histological subtypes. A strong positive signal for PTCH2 mRNA was observed exclusively in the tumor cells of all BCCs. Notably, the signal was consistently stronger in the palisading peripheral cells of the tumor nests (Fig. 2 ). These cells also showed a positive immunostaining for the cell proliferation marker, Ki-67.

The finding that in BCCs having frequent mutations in the PTCH1 gene, the expression of the PTCH2 mRNAs is upregulated, tightly links the novel PTCH2 according to the invention with the PTCH/SHH cascade of signalling events. It is therefore likely that PTCH2 represents a target gene of this pathway which is under the negative regulation of PTCH1, precisely as PTCH1 itself. Moreover this observation strongly suggests that PTCH2 has functions distinct from PTCH1 since upregulation of PTCH2 expression appears unable to compensate for inactive PTCH1 protein. This conclusion is also supported by the early embryonic lethality seen in PTCH1 (-/-) mice 5,13) and the lack of genetic heterogeneity in Gorlin syndrome. However, whether PTCH2 may block the constitutive signalling of SMO, or could act as an additional SHH receptor, possible dependent on alternative splicing, remains as the subject of further experimentation.

CLAIMS

1. An isolated human protein or an analogue or variant thereof capable of participating in the human PTCH/SHH pathway during embryonic development and/or carcinogenesis comprising at least about 1040 amino acids of SEQ ID NO: 1.
- 5 2. A protein according to claim 1, which is essentially comprised of SEQ ID NO: 1.
3. A nucleic acid encoding a protein according to any one of claims 1 and 2.
4. A nucleic acid encoding a protein according to claim 3 comprising at least about 3094 bases of SEQ ID NO: 2.
- 10 5. An isolated genomic nucleic acid comprising parts or all of SEQ ID NO: 5.
6. An isolated variant of the nucleic acid according to claim 5.
7. A nucleic acid having the sequence of a splicing variant selected from the group of variants defined in Figure 2B and SEQ ID NO:3 and SEQ ID NO:4.
8. An isolated nucleic acid capable of specifically hybridising to a nucleic acid according to any one of claims 3-7.
- 15 9. A protein or polypeptide encoded by a nucleic acid according to claim 7.
10. A protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9, for use as a medicament.
11. Use of a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 in the manufacture of a medicament for the treatment of a condition involving tumors, such as BSS.
- 20 12. A method of in vitro or in vivo diagnosis, wherein a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 is used.
13. A method of screening wherein a library of suitable candidate compounds is screened for modified drugs using a protein according to claim 1 or 2 or a polypeptide according to claim 9 as a lead compound.
- 25 14. A method of synthesis of a modified drug, wherein a protein according to claim 1 or 2, or a protein or polypeptide according to claim 9, is used.
15. A modified drug identified by the method according to claim 13 or synthesized according to claim 14.

16. A vector comprising a nucleic acid according to any one of claim 3-8.
17. A recombinant cell comprising a vector according to claim 16.
18. An antibody which specifically binds to a protein or polypeptide according to claim 1, 2 or 9.
- 5      19. A recombinant cell expressing an antibody according to claim 18.
20. A kit for the detection of a human PTCH2 gene or polypeptide comprising in a container a molecule selected from the group consisting of a nucleic acid according to any one of claims 3-8, a polypeptide or protein according to claim 1, 2 or 9 or an antibody according to claim 18.
- 10     21. Use of a nucleic acid according to any one of claims 3-7 in gene therapy.
22. Use of nucleic acid according to any one of claims 3-7 as a probe, a primer or a diagnostic reagent.

15

The intron sequences between exons 2 - 3 and exons 18 - 19 are missing (introns: small letters, exons: capital letters). Small letters in the first exon indicate nucleotides that have not been unambiguously determined.

*Exon 1*

1 CGGGTGAATC CCGGCGCCGC GCCCCGGACC CGCAGCTCCC TGCACTCCTC  
51 CCTCCCAGCC GCTTTAACAC CCACACCCCCA CAGTCTCTCC CACGsCCGCG  
101 CCTTGGCGGC CCCACTGAAT CCCTACGCGG GGCCCAGCGG TACCGGGAGA  
151 CCGGGCTAGC CTATGGGAGC GCCCAGATAA CGCGGGTTGG GGGCGCCCGC  
201 GCCCcCATCC CCGCCAG**C A T** GACTCGATCG CCGCCCCTCA GAGAGCTGCC  
251 CCCGAGTTAC ACACCCCCAG CTCGAACCGC AGCACCCAG gtgagtagag  
301 ggggagctgg aagaaggaag agagcggagc caggtctgtc actcgccct  
351 ctgcaagggtt tgtatgtct tgaagtgccg agtgtcatta gatgtctgaa  
401 ggcaagttag agccagcacc gcaagcaagt tgtgcgtgtg tgtcggtgtg  
451 tctgtgccgg tgtctcctca tcgtctggcc agtgagaatg aatgtctgtg  
501 gttcacctc tgtgtccacc cgacgacagg tgtgtgtaca tatgtatcct  
551 gctctcagaa aatgggccta tgccgccggg cgccgtgact cacgcctgta  
601 atcccaacac tgggaggctg aggcaaggcag attacctgag gtcaggagtt  
651 cgagaccagc caggccaaca tggggaaact ctgtctctac taaaaataaa  
701 aattagcagg gcgtggtggc gggcgctgt agtcccaact actcgaggagg  
751 ctgaggcagg agaatctctt gaacctggga ggcggaggtt gcagtcaagc  
801 cgagatcaca ccactgcact ccagccaggg caacagagcg agatcgct  
851 caaaaaaaaaaaa aaaaaaaaaaaa aaaaggagag aaaacaaaaaa gaaaagaaaag  
901 gaaaataggc ctatgccttc ctcaggtgtg tgctgggat ggtgggtgtt  
951 acatcttcca agtctggcc tgtgtctgtg ttggtgctcc ctgtcccaca  
1001 tccagaaaatc aagaagcgag ggctggcag cagatataca gggtgagaag

Fig. 1

2/13

1051 ggaaggattt catgcattgt tacagtgatg cctggctgac ctttccttt  
EXON 2  
1101 ccatcccagA TCCTAGCTGG GAGCCTGAAG GCTCCACTCT GGCTTCGTGC  
1151 TTACTTCCAG GGCCTGCTCT TCTCTCTGGG ATGCAGGGATC CAGAGACATT  
1201 GTGGCAAAGT GCTCTTCTG GGACTGTTGG CCTTTGGGGC CCTGGCATTA  
1251 GGTCTCCGCA TGGCCATTAT TGAGACAAAC TTGGAACAGC TCTGGGTAGA  
1301 AGTGGGCAGC CGGGTGAGCC AGGAGCTGCA TTACACCAAG GAGAAGCTGG  
1351 GGGAGGAGGC TGCATACACC TCTCAGATGC TGATACAGAC CGCACGCCAG  
1401 GAGGGAGAGA ACATCCTCAC ACCCGAAGCA CTTGGCCTCC ACCTCCAGGC  
1451 AGCCCTCACT GCCAGTAAAG TCCAAGTATC ACTCTATGGG AAG.....  
1501 ..... g  
1551 tgagtctggc tgagccctg agcagctggg ggcgaggcgt gctgtggggg  
1601 ttctggagtg ggaatcccct tcttctgctg atctcctatg cccctggcta  
EXON 4  
1651 ttgcagTCCT GGGATTGAA CAAAATCTGC TACAAGTCAG GAGTTCCCCT  
1701 TATTGAAAAT GGAATGATTG AGCGGgttaag tgtcctgaga gggagtagag  
1751 gcagaacttt ttctgttagcg tgggaggact cagagaccga gcaagcccc  
1801 cagcctgcaa tctgccccct taaaactaag gagggggatt gcagagggca  
1851 tcctacaaag gttgtggggc aggactgacg tggcccgaaaa tatccctggc  
EXON 5  
1901 agATGATTGA GAAGCTGTTT CCGTGCGTGA TCCTCACCCCC CCTCGACTGC  
1951 TTCTGGGAGG GAGCCAAACT CCAAGGGGGC TCCGCCTACC TGCCgtgagt  
2001 gccactcctg gggccctgct tcatactcccg ctggggactc tcccaagcaga  
2051 aaggaggggt ctggggaatg aggatgatca aaaccttacc aaggtcctaa  
2101 ttacctccca ggccaggaac agagagcatg ggcttccca aggctctctc  
2151 cacatcctcc ttctctttcc ctctcaagga aggaagacct gacttattta  
2201 cacaaaacta aacacaaaaga tctgttaagat ctgagcaaag gagaaaaaga  
2251 tccccacaaa gaggcttgc tggggaaaat taccttaggtg tttgctaagc  
2301 cattgcccag gccagaaaaga aaacctgcta caggcatgtg cctgctggtt  
2351 gtatattaga accaagcaca cagcttgta aggaactcag tggggccttt

Fig. 1 (cont.)

3/13

2401 ctggccctt tctatgtatt aggttaaccct gccctgatat tcgttcagc  
2451 cccttgtact cttctacagc tcactgttagc accctggtgg gcccattgcag  
2501 cctggcagtt ctgagaagct gaggcttgca caccctccat atgaaaggac  
2551 aaatcgccag ataagaggag ggtgggtac agcatggcgc cccagcagca  
2601 gtttggagcc tgggtttcg tccctgaccc tcaccaacta taggctttc  
EXON 6  
2651 cctcagCGGC CGCCCGGATA TCCAGTGGAC CAACCTGGAT CCAGAGCAGC  
2701 TGCTGGAGGA GCTGGGTCCC TTTGCCTCCC TTGAGGGCTT CCGGGAGCTG  
2751 CTAGACAAGG CACAGGTGGG CCAGGCCTAC GTGGGGCGGC CCTGTCTGCA  
2801 CCCTGATGAC CTCCACTGCC CACCTAGTGC CCCCAACCAT CACAGCAGGC  
2851 AGgtgggttc caaccaggc tgccaggaa aggctgttt cttcccttt  
2901 cccttcctca tactcctgtg ttctggggaa gctgactgct ctgtgccctg  
2951 accccccaact tcctggccat tattaccctg ctcccacagt gccaggcccc  
3001 caatgttcca ttcccattca gttatcctac ggagccctca agtggatat  
3051 atgaatccct tttcccttt ctaagcctag ataaggctgg acttctttt  
3101 tttttttttt ttgagtcctca ctctgtcacc caggctggag tgcagtagtt  
3151 cgatcttggc tcactgcaac ctcggctcaa gcaattctcc tgccttagcc  
3201 tcctgagtag ctgggattac aggtgcccac caccatgccc ggctaatttt  
3251 tattagcctc ccaaagtgct gggattacag gcgtgagcca ctgcgcctgg  
3301 ccaaggctgg acttttatac aaaatagact aatacaggaa aactaagaac  
3351 acagcaggta agcatgaata tcatacctgg tttccctagg tttttgtgg  
3401 ccctgcaaatt gtggacttt tttcagaatc cgccagttac accagctcct  
3451 cccagaagcc tacttccagg cctctgcttc cccttggggc ttccctgtctg  
3501 cgggatacta gctgttcaact cctgcagagc agtcaagagg ctcagaataag  
3551 ttacctacac tccagcccta ctgagcttca tggcagcgtg gttcctggag  
3601 gtggaagccc agggacactc agttatccac ggccagggcc ttgagcatta  
EXON 7  
3651 acccctcctg ttcccctcca gGGCTCCCAA TGTGGCTCAC GAGCTGAGTG  
3701 GGGGCTGCCA TGGCTTCTCC CACAAATTCA TGCACGGCA GGAGGAATTG

Fig. 1 (cont.)

4/13

3751 CTGCTGGAG GCATGGCCAG AGACCCCCAA GGAGAGCTGC TGAGGttaggg  
3801 ttcctctgg gagtttgtga ggggactctg ttcatgagaa cccatactgt  
3851 aatgccaggc agctctggca aaaggccctt cacatccctc accaggtgtt  
EXON 8  
3901 tgggcagct ctgacccttg gttctccac acccccacca gGGCAGAGGC  
3951 CCTGCAGAGC ACCTTCTTGC TGATGAGTCC CCGCCAGCTG TACGAGCATT  
4001 TCCGGGTGA CTATCAGACA CATGACATTG GCTGGAGTGA GGAGCAGGCC  
4051 AGCACAGTGC TACAAGCCTG GCAGCGCGC TTTGTGCAGg tcggtatgga  
4101 caaggacaag ggggtgtccc tgaggccatt ccctcctcct gccccctcct  
EXON 9  
4151 atccaccctg tttctccagC TGGCCCAGGA GCCCCTGCCT GAGAACGCTT  
4201 CCCAGCAGAT CCATGCCTTC TCCTCCACCA CCCTGGATGA CATCCTGCAT  
4251 GCGTTCTCTG AAGTCAGTGC TGCCCGTGTG GTGGGAGGCT ATCTGCTCAT  
4301 Ggtgggtctt gcacctggca cttggccccc accccaccc caaccagtgc  
EXON 10  
4351 ccaccctggg agccctgag actgccctt ccccccacag CTGGCCTATG  
4401 CCTGTGTGAC CATGCTGCGG TGGGACTGCG CCCAGTCCCA GGGTTCCGTG  
4451 GGCCTTGCCG GGGTACTGCT GGTGGCCCTG GCGGTGGCCT CAGGCCTTGG  
4501 GCTCTGTGCC CTGCTCGGCA TCACCTCAA TGCTGCCACT ACCCAGgtac  
4551 gccaggactg cagggcagac tcagtgccag tcaccaggct tcacgggtcc  
EXON 11  
4601 tcagctgccc gtcctctgc ccctccagGT GCTGCCCTTC TTGGCTCTGG  
4651 GAATCGGCGT GGATGACGTA TTCCTGCTGG CGCATGCCTT CACAGAGGCT  
4701 CTGCCTGGCA CCCCTCTCCA Ggtggggcct tgtccccag ggctcatctg  
4751 aggcaactca gcttactggtaa agagccctc ttggttcaag tgacccttgg  
4801 gctgctaattg aacctcggtg cctttgtcc ccatctgtaa acaggggaaa  
4851 taatagtgtctgtc gtttactggtaa ggttattgtt tggatcagtg aggttaactca  
4901 agttgaatgc tttagaacagc ccatcatacg tacatggtac ccaataaaatg  
4951 ctagccactg tgttatgact gccccaccc tcaccccaa gttcctgagc  
5001 ctcccccttca ctccactttg acacggcccc tcccttgta cctgagggca  
EXON 12  
5051 ggtccccact ctgtcctggc agGAGCGCAT GGGCGAGTGT CTGCAGCGCA

Fig. 1 (cont.)

5/13

5101 CGGGCACCAAG TGTGCTACTC ACATCCATCA ACAACATGGC CGCCTTCCTC  
5151 ATGGCTGCCCG TCGTTCCCCT CCCTGCCTG CGAGCCTTCT CCTACAGGC  
5201 GGCCATAGTG GTTGGCTGCA CCTTTGTAGC CGTGATGCTT GTCTTCCCAG  
5251 CCATCCTCAG CCTGGACCTA CGGGGGCGCC ACTGCCAGCG CCTTGATGTG  
5301 CTCTGCTGCT TCTCCAGgt a ctgcgtgcgc cccagccccct tcctcccgtg  
5351 acccaacgcca gcctgtcccc tcaccagcat ttcaaggcac agacctgtca  
EXON 13  
5401 tccactctct acctcttcca gTCCCTGCTC TGCTCAGGTG ATTCAAGATCC  
5451 TGCCCCAGGA GCTGGGGGAC GGGACAGTAC CAGTGGGCAT TGCCCACCTC  
5501 ACTGCCACAG TTCAAGCCTT TACCCACTGT GAAGCCAGCA GCCAGCATGT  
5551 GGTCACCATC CTGCCTCCCC AAGCCCCACCT GGTGCCCCCCA CCTTCTGACC  
5601 CACTGGGCTC TGAGCTCTTC AGCCCTGGAG GGTCCACACG GGACCTTCTA  
5651 GGCCAGGAGG AGGAGACAAG GCAGAAGGCA GCCTGCAAGT CCCTGCCCTG  
5701 TGCCCCGCTGG AATCTTGCCC ATTCGCCCCG CTATCAGTTT GCCCCGTTGC  
5751 TGCTCCAGTC ACATGCTAAG gtaagactgg gcagagcagg gcagagactt  
5801 agcatctctg ggcccagaag ggcagagagg gcttagtcca ctgcctgagg  
EX  
5851 ggctgggggc agccctgggg tctccagctt agttgctaca tcccgca gGC  
XON 14  
5901 CATCGTGCTG GTGCTCTTG GTGCTCTTCT GGGCCTGAGC CTCTACGGAG  
5951 CCACCTTGGT GCAAGACGGC CTGGCCCTGA CGGATGTGGT GCCTCGGGGC  
6001 ACCAAGGAGC ATGCCTTCCT GAGCGCCAG CTCAGGTACT TCTCCCTGTA  
6051 CGAGGTGGCC CTGGTGACCC AGGGTGGCTT TGACTACGCC CACTCCAAAC  
6101 GCGCCCTCTT TGATCTGCAC CAGCGCTTCA GTTCCCTCAA GGCGGTGCTG  
6151 CCCCCACCGG CCACCCAGGC ACCCCGCACC TGGCTGCACT ATTACCGCAA  
6201 CTGGCTACAG Ggtgagagggc gagggagacgg gcagggaggg gtgctgcagg  
6251 gagaaacgcc ctggggccac cagctaata g aaccctatcc tggctccccc  
EXON 15  
6301 cagGAATCCA GGCTGCCCTT GACCAGGACT GGGCTTCTGG GCGCATCACC  
6351 CGCCACTCGA CCGCAATGGC TCTGAGGATG GGGCCCTGGC CTACAAGCTG  
6401 CTCATCCAGA CTGGAGACGC CCAGGAGCTT CTGGATTCA GCCAGgttgg

Fig. 1 (cont.)

6/13

6451 gagagggctg gagggtcca ctagtacagg ggctgcaggc ctccctggcc  
EXON 16  
6501 caggcttca gccctctctg cctctgcagC TGACCACAAG GAAGCTGGTG  
6551 GACAGAGAGG GACTGATTCC ACCCGAGCTC TTCTACATGG GGCTGACCGT  
6601 GTGGGTGAGC AGTGACCCCC TGGGTCTGGC AGCCTCACAG GCCAACTTCT  
6651 ACCCCCCACC TCCTGAATGG CTGCACGACA AATACGACAC CACGGGGAG  
6701 AACTTCGCA gtgagtcttg gggggagctc ggcaagagcc tcagcctcgc  
6751 ccacacaagc cctgagcctg aggcctgcc cactctgccc cgtgctcacc  
EXON 17  
6801 gccctgtccc tctccctctt ctcccttccc ctcccttcca cagTCCCAGCC  
6851 AGCTCAGCCC TTGGAGTTTG CCCAGTTCCC TTTCTGCTG CGTGGCCTCC  
6901 AGAAGACTGC AGACTTTGTG GAGGCCATCG AGGGGGCCCG GGCAGCATGC  
6951 GCAGAGGCCG GCCAGGCTGG GGTGCACGCC TACCCCAGCG GCTCCCCCTT  
7001 CCTCTTCTGG GAACAGTATC TGGGCCTGCG GCGCTGCTTC CTGCTGGCCG  
7051 TCTGCATCCT GCTGGTGTGC ACTTTCCTCG TCTGTGCTCT GCTGCTCCTC  
7101 AACCCCTGGA CGGCTGGCCT CATAgtgagt gcttgcagga gtggggacag  
7151 agacacccca cccttccctg cccagcctgt catccctcct gccaggagcc  
EXON 18  
7201 ctctgtgagc cctgtctccc tcagGTGCTG GTCCTGGCGA TGATGACAGT  
7251 GGAACTCTTT GGTATCATGG GTTTCCTGGG CATCAAGCTG AGTGCATCC  
7301 CCGTGGTGAT CCTTGTGGCC TCTGTAGGCA TTGGCGTTGA GTTCACAGTC  
7351 CACGTGGCTC TGGGCTTCCT GACCACCCAG GGCAGCCCGA ACCTGCGGGC  
7401 CGCCCATGCC CTTGAGCACA CATTGCCCC CGTGACCGAT GGGGCCATCT  
7451 CCACATTGCT GGGTCTGCTC ATGCTTGCTG GTTCCCACCT TGACTTCATT  
7501 GTAAG.....  
7551 ..... gtagggaggg ctcggggcag ggaggcaggg ctcaggacag  
EXON 20  
7601 gcctgggctg actccccca caccctaccc ctagGTACTT CTTTGCAGCG  
7651 CTGACAGTGC TCACGCTCCT GGGCCTCCTC CATGGACTCG TGCTGCTGCC  
7701 TGTGCTGCTG TCCATCCTGG GCCCGCCGCC AGAGgtgacc acaccctcg  
7751 caccatccct ctactccct cccaaaggac ggggttaggga gaggcaagg

Fig. 1 (cont.)

7/13

7801 aaggacaga gccctgtggc ccacagacag gtacccccc aacaggtgcc  
7851 accagctgaa ggtggcagcc tcctccttc cccagacacc atgttcctgc  
7901 ccctcagccc tcctggcttc ttcatggac ccaccttaga cttaggat  
7951 ccagaacaag gtgcagggtt tgccccaggc ctaaacatcc tgcgcctgc  
8001 cagctctcat atcctgctgg agaccaacaa gggccccagc ttcccaacag  
8051 tcatggtaat ccccgagcgag atgctaaagg ggacgggagc cccaggggcc  
EXON 21  
8101 cgtgggctta ctggggctgg tgtctccca cagGTGATAc AGATGTACAA  
8151 GGAAAGCCCA GAGATCCTGA GTCCACCAGC TCCACAGGGGA GGCAGGGCTTA  
8201 Ggtggggggc atcctcctcc ctgccccaga gctttgcac agtactacc  
8251 tccatgaccg tggccatcca cccacccccc ctgcctggc cctacatcca  
8301 tccagccccct gatgagcccc cttggcccccc tgctgtcact agctctggca  
8351 acctcagttc caggggacca ggtccagcca ctgggtgaaa gagcagctga  
8401 agcacagaga ccatgtgtgg ggcgtgtggg gtcactggga agcactgggt  
8451 ctggtgtag acgcaggatg gaccctgga gggctctgct gctgctgcat  
8501 cccctctccc gacccagctg tcatggcct ccctgatatac catacagaac  
8551 agccaccgat ttgcacatcc aggccctgtgt gagcctgtat ctgtgtcact  
8601 tgagagtgaa agctggcact tggggctgca gtgcagccct gtcccccttc  
8651 ccacccccaca ccactgcctg cccagctgac caagcctgag ggaccctcca  
8701 gcacccttcc gtctggtgac tcctggcag gctctccata tccctgcca  
8751 cctcctacca catccattat ttatatgaaa atgtctattt ttgtgtata  
8801 catacatgtt agctatgtat aaagtttat tttttaaaga atgaaatata  
8851 ttctatgtga agctatgtat aaagtttat tttttaaaga atgaaatata  
8901 ttctatgtga actaatctcg aaagtttat tttttaaaga atgaaatata  
8951 ttctatgtgt gcaagtgaac attagcttca gttgctttt tttggacaga  
9001 gtggggagtt tgcaagtgaa cattagctat tggaaaggagc ttctctggg  
9051 ccaggacctg aggtattagc ttctcttagtt ctgggtggaa aagacccca  
9101 attctggatt ttgtcatat acttgtaac atcatctgga ttaagtgcct

Fig. 1 (cont.)

8/13

9151 actatacataaa acgataaacaat attttgttgg tgtgaaatcc tactgggttc  
9201 aatctggaga ccgagagcag aaaaaaaaaaaga acccccactgt gtggctttca  
9251 gagccaccat attccagcct gcccgctct ccagactcac ctccacacctac  
9301 ctgcttcacc cgcacggaa acggcaaggc agaggggcaa agccatgcag  
9351 caggttggaaag gcgaggttggaa ggcagatcag gaaagcagcc agttgaagca  
9401 gagagaggc aacagggtct ggggagcttc tcaggagggtt tgtggacc  
9451 gggaaaggag ccaggttcca gagcaacctc caaggcaaag gcctctgtaa  
9501 gttgggttgc tcgacagccg agaggtgtct ttggccagtc agccagtgg  
9551 tcagttgcgg gaactgctca gaaaactgagg tgctagcagt tagtgaggac  
9601 acagcgtaag ttgtttgttc tgtgaaagtt gaacagctcc actaagcaga  
9651 ggccttgaag agtggccaca gcccttgaat agagcacaga gcctcaccta  
9701 gaggttgcgg gaggttgca actgccccctt cccagccata gcttaggacc  
9751 catagtctag ttcacataga ccctgggctc caaccaccca ctcaccagga  
9801 atgatcccac cccaggaaca atgcgttctc acatcccacc ccacctggac  
9851 aaaggccagg aaatcatgtt ctgacaaaaa gatacaacaa caaaaacaac  
9901 aacaacaaaa aacgcctatt gcaattgaat ccacgctaaa atgcctaaaa  
9951 agctcaagag aagcgggttag ttggcagaga acctagagta gggggtgcaa  
10001 ccagcaggcc caagggaggg aggctgcatt tgggtccagc agtgtttgg  
10051 tcaccaagaa gggccttcta ggtggagcag agagagctca ccaggccaga  
10101 atagtgcaaa gggggtcagc cctcagtgcc acttaccagc ggagtaaccc  
10151 tggcaagtt agccagcctc actaagcctc cccatctca tctttccagG  
XON 22 E  
10201 CCCGAGGAGA TCTAGCCTCT GCCTCCCCACC CCAGCACCCCC CTCATCAGAC  
10251 ACAAGGGAGCG CCACTGTCTG GACAGGGCTGA ATTGGTCTTC GGGTCCCTAA  
10301 TTTCTCATAC GCCATTCCCT CTGCCTAGAA CACTTTCTCA CCTCCCCTTG  
10351 ATGTGACCCCC ATATCACCCCT TCGAGGGTGAA TTGGATCGGA TGCCATCTCC  
10401 TCCAGGAGGG GTGGGGTCGT GCCTCCTGTG AGGTCCCAGT GCCCCCTGAGT  
10451 GTCTGTGCCCT GTCTGTTCC CCGTCCCTCT CTCTAACGCC GGAGGCTTAC

Fig. 1 (cont.)

9/13

10501 TGCAGGGTAAG GACGGCGGGA CAGGACCTTA ACCCCTGGGA CGAACACCAG  
10551 CTCCGCAAAG GACTCCGCAC CCGGCGCCGC CCACGGGGTG CGGGTCCCAG  
10601 GAGGACCAGC AGAGAGGAGC ATAGGAGAGC AAAGGAGATC AGTACCCAT  
10651 GGCTTCCCCG GTGGCGCGGA ACAGCCCGGA GCCGCCTGTG ATTTGCATAAC  
10701 CCATGGTGCA CCACGAAAAG ATACCCCTCAA GATGCTTGCA CTCCCTCTGT  
10751 GCGCGCATTT CTGCACTGTT TTAGAGCATG ATGCCTCTTA CACGCATCTG  
10801 TGTGCATAAA CTACATATAG GGAGTGCAGTA CCACGCAGGC ATCCAACAAC  
10851 CATAAGTGTG TTAAGTGTAA GTTCTCCCTG CGAGGTTCGA AGCGGAAGTC  
10901 ACGAATATAC TCAGGTTTCT CTTCAAAGCG CATAAAATCTT TCGCCTTTA  
10951 CTAAAGATTT CCGTGGAGAG AAAGTTGTGA GTTTTTATTG AATTTTTGAA  
11001 GGCCTCTTAT TTCCTGAGGC TACATTTTA AGTATTAAAA GTTAGGCAAC  
11051 TACAAAAAAA AAAAAAAA

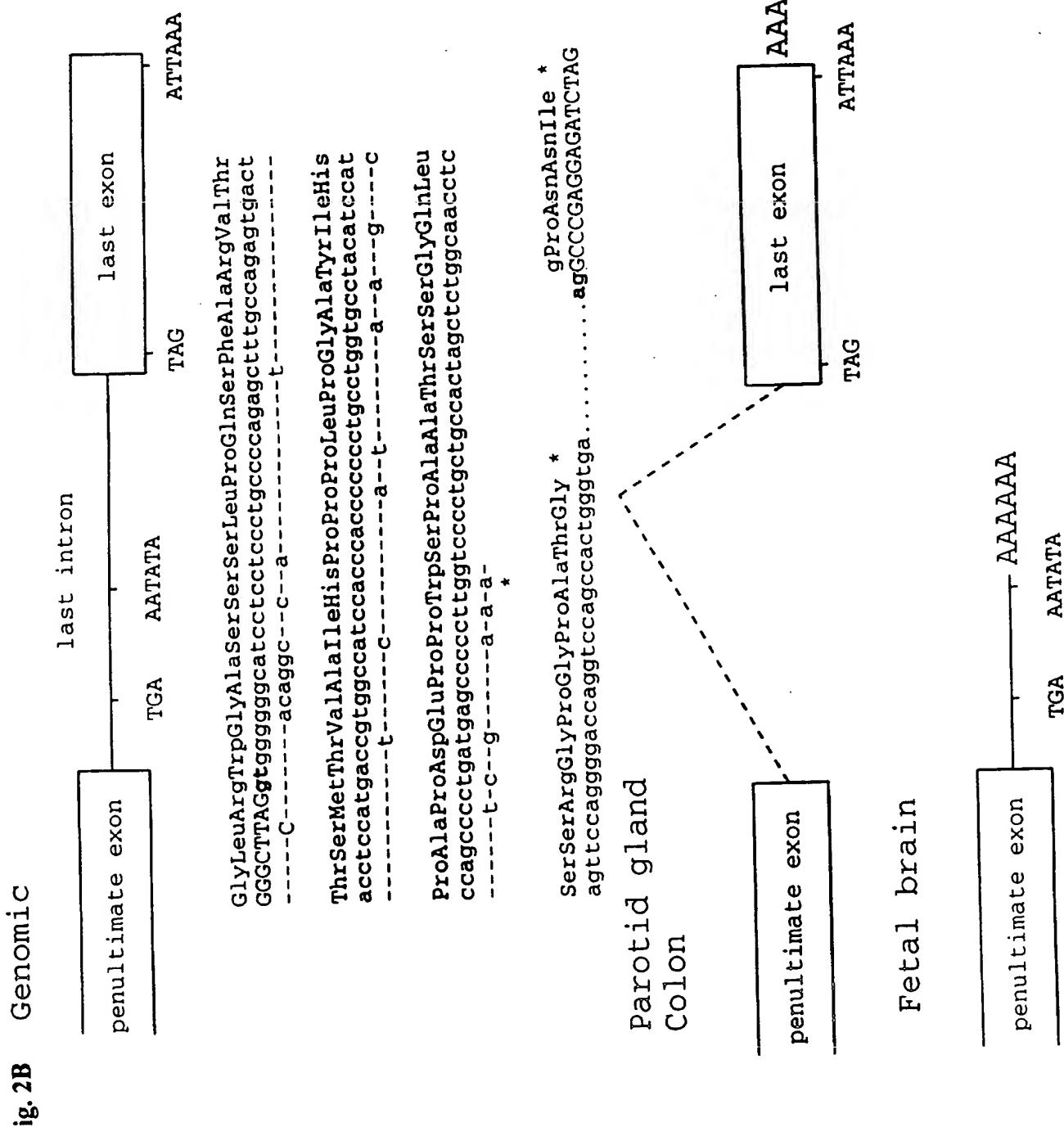
**Fig. 1 (cont.)**

10/13

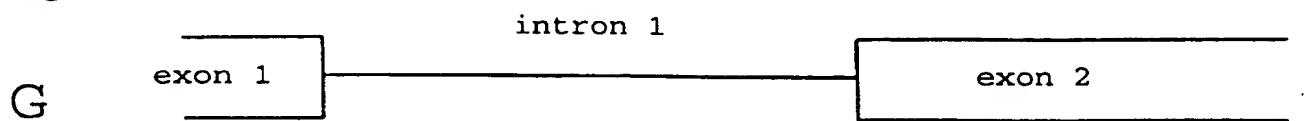
1 .....MTRSPPLRELP..... 11  
 1 MASAGNAEAEQDRGGGGSGCIGAPGRPAGGGRRRTGGLRRAAAPDRDYL 50  
 12 ..PSYTPPARTAAPQI...LAGSLKAPLWLRAYFQGLLFSLGCGIQRHCG 56  
 51 HRPSYCDAA.FALEQISKGKATGRKAPLWLRAKFQRLLFKLGCYIQKNCG 99  
 57 KVLFLGLLAFGALALGLRMAIIETNLEQLWVEVGSRVSQELHYTKEKLGE 106  
 100 KFLVVGLLIFGAFAVGLKAANLETNVEELWVEVGRVSRELNYTRQKIGE 149  
 107 EAAYTSQMLIQTARQEGENILTPPEALGLHLQAALTASKVQVSLYGKSDL 156  
 150 EAMFNPQLMIQTPKEEGANVLTTEALLQHLDSDLQASRVHVYMYNRQWKL 199  
 157 NKICYKSGVPLIENGMIERMIEKLFPCVILTPLDCFWEAGAKLQGGSAYLP 206  
 200 EHLCYKSGELITEGYMDQIEYLYPCLIITPLDCFWEAGAKLQSGTAYLL 249  
 207 GRPDIQWTNLDPQLLEELGPFA.SLEGFRELLDKAQVGQAYVGRPCLHP 255  
 250 GKPPPLRWTNFDPLEFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPCLN 299  
 256 DDLHCPPSAPNHHSRQAPNVAHELSSGGCHGFSHKFMHWQEELLGGMARD 305  
 300 ADPDCPATAPNKNSTKPLDMALVLNGGCHGLSRKYMHWQEELIVGGTVKN 349  
 306 PQGELLRAEALQSTFLMSPRQLYEHFRGDYQTHDIGSEEQASTVLOAW 355  
 350 STGKLVSAHALQTMFQLMTPKQMYEHFKGYEYVSHINWNEDKAAAILEAW 399  
 356 QRRFVQLAQEALPENASQOIHAFSSTLDDILHAFSEVSAARVVGYYLLM 405  
 400 QRTYVEVVHQSVAQNSTQKVLSTTTLDDILKSFSDVSVIKVASGYLLM 449  
 406 LAYACVTMLRWDCQAQSQGSVGLAGVLLVALAVASGLGLCALLGITFNAAT 455  
 450 LAYACLTMLRWDCSKSQGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAT 499  
 456 TQLPFLALGIGVDDVLLAHAFTEALPG..TPLQERMGECLQRTGTSVV 503  
 500 TQLPFLALGVGVDDVLLAHAFSETGQNKRIPFEDRTGECLKRTGASVA 549  
 504 LTSINNMAAFLMAALVPIPALRAFLQAAIVVGCTFVAVMLVFPAILSLD 553  
 550 LTSISNVTAFFMAALIPIPALRAFLQAAVVVVFNFAMVLLIFPAILSMD 599  
 554 LRRRHQCRLDVLCCFSSPCSAQVIQILPQELGDGT.....VPVG 592  
 600 LYRREDRRLDIFCCFTSPCVSRVIQVEPQAYTDHDNTRYSPPPPSSHS 649  
 593 IAH....LTATVQAFTHCEASSQHVVTILPPQAH....VPPPSDPLGS 633  
 650 FAHETQITMQSTVQLRTEYDPHTHYYTTAEPRSEISVQPVTVTQDTLSC 699

Fig. 2A

11/13



12/13

**Fig. 2C**

Pro Gln  
CCC CAG gt.....gt.....ag ATC CTA.....Cag GGC  
← 67 bp → 755 bp → ← 51 bp →



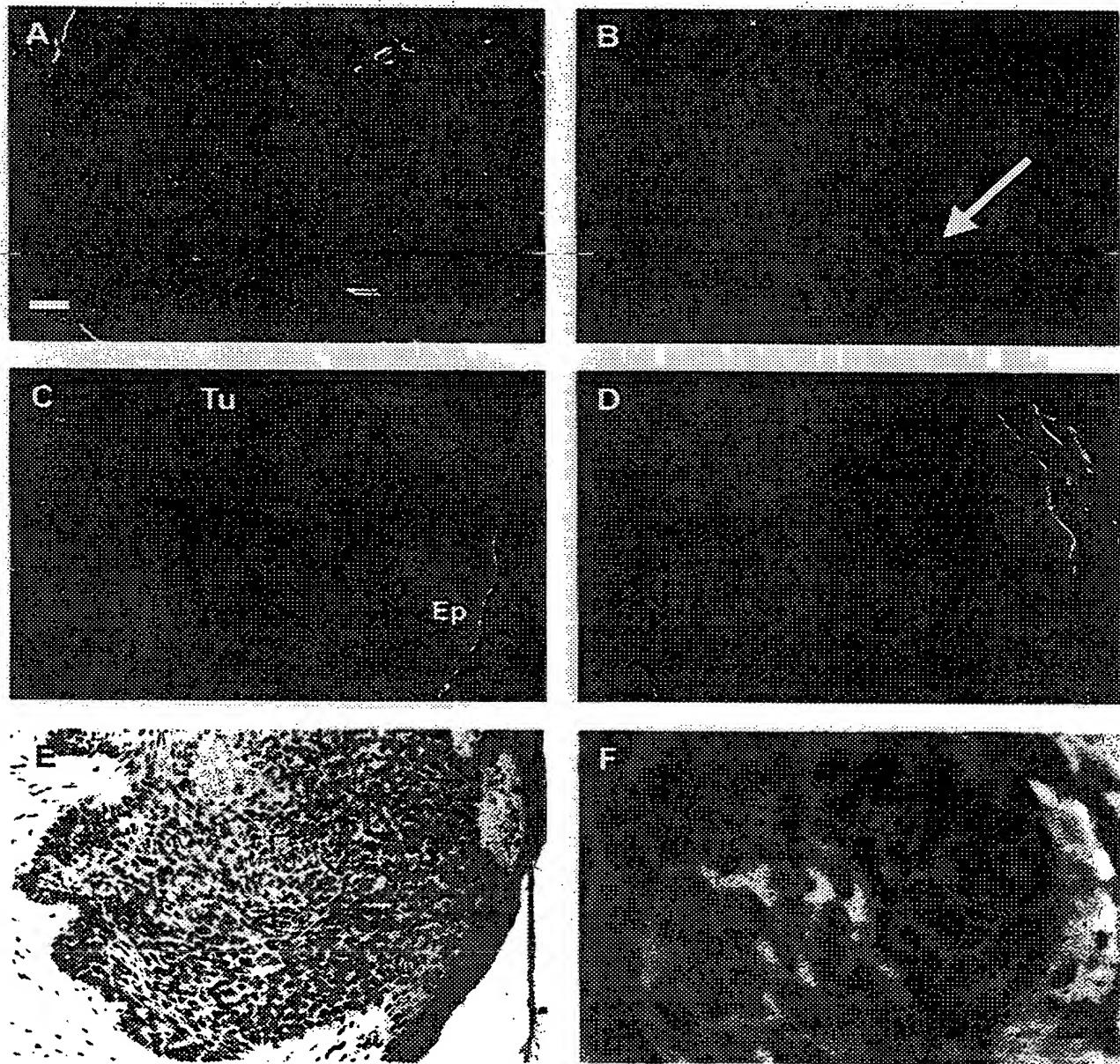


Fig. 3

## SEQUENCE LISTING

<110> PHARMACIA & UPJOHN S.P.A.  
KAROLINSKA INNOVATIONS AB

<120> A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY

<130> 56822

<140>  
<141>

<160> 5

<170> PatentIn Ver. 2.1

<210> 1  
<211> 1146  
<212> PRT  
<213> HUMAN

<400> 1  
Met Thr Arg Ser Pro Pro Leu Arg Glu Leu Pro Pro Ser Tyr Thr Pro  
1 5 10 15  
Pro Ala Arg Thr Ala Ala Pro Gln Ile Leu Ala Gly Ser Leu Lys Ala  
20 25 30  
Pro Leu Trp Leu Arg Ala Tyr Phe Gln Gly Leu Leu Phe Ser Leu Gly  
35 40 45  
Cys Gly Ile Gln Arg His Cys Gly Lys Val Leu Phe Leu Gly Leu Leu  
50 55 60  
Ala Phe Gly Ala Leu Ala Leu Gly Leu Arg Met Ala Ile Ile Glu Thr  
65 70 75 80  
Asn Leu Glu Gln Leu Trp Val Glu Val Gly Ser Arg Val Ser Gln Glu  
85 90 95  
Leu His Tyr Thr Lys Glu Lys Leu Gly Glu Glu Ala Ala Tyr Thr Ser  
100 105 110  
Gln Met Leu Ile Gln Thr Ala Arg Gln Glu Gly Glu Asn Ile Leu Thr  
115 120 125  
Pro Glu Ala Leu Gly Leu His Leu Gln Ala Ala Leu Thr Ala Ser Lys  
130 135 140  
Val Gln Val Ser Leu Tyr Gly Lys Ser Trp Asp Leu Asn Lys Ile Cys  
145 150 155 160  
Tyr Lys Ser Gly Val Pro Leu Ile Glu Asn Gly Met Ile Glu Arg Met  
165 170 175  
Ile Glu Lys Leu Phe Pro Cys Val Ile Leu Thr Pro Leu Asp Cys Phe  
180 185 190  
Trp Glu Gly Ala Lys Leu Gln Gly Gly Ser Ala Tyr Leu Pro Gly Arg  
195 200 205

Pro Asp Ile Gln Trp Thr Asn Leu Asp Pro Glu Gln Leu Leu Glu Glu  
210 215 220

Leu Gly Pro Phe Ala Ser Leu Glu Gly Phe Arg Glu Leu Leu Asp Lys  
225 230 235 240

Ala Gln Val Gly Gln Ala Tyr Val Gly Arg Pro Cys Leu His Pro Asp  
245 250 255

Asp Leu His Cys Pro Pro Ser Ala Pro Asn His His Ser Arg Gln Ala  
260 265 270

Pro Asn Val Ala His Glu Leu Ser Gly Gly Cys His Gly Phe Ser His  
275 280 285

Lys Phe Met His Trp Gln Glu Glu Leu Leu Gly Gly Met Ala Arg  
290 295 300

Asp Pro Gln Gly Glu Leu Leu Arg Ala Glu Ala Leu Gln Ser Thr Phe  
305 310 315 320

Leu Leu Met Ser Pro Arg Gln Leu Tyr Glu His Phe Arg Gly Asp Tyr  
325 330 335

Gln Thr His Asp Ile Gly Trp Ser Glu Glu Gln Ala Ser Thr Val Leu  
340 345 350

Gln Ala Trp Gln Arg Arg Phe Val Gln Leu Ala Gln Glu Ala Leu Pro  
355 360 365

Glu Asn Ala Ser Gln Gln Ile His Ala Phe Ser Ser Thr Thr Leu Asp  
370 375 380

Asp Ile Leu His Ala Phe Ser Glu Val Ser Ala Ala Arg Val Val Gly  
385 390 395 400

Gly Tyr Leu Leu Met Leu Ala Tyr Ala Cys Val Thr Met Leu Arg Trp  
405 410 415

Asp Cys Ala Gln Ser Gln Gly Ser Val Gly Leu Ala Gly Val Leu Leu  
420 425 430

Val Ala Leu Ala Val Ala Ser Gly Leu Gly Leu Cys Ala Leu Leu Gly  
435 440 445

Ile Thr Phe Asn Ala Ala Thr Thr Gln Val Leu Pro Phe Leu Ala Leu  
450 455 460

Gly Ile Gly Val Asp Asp Val Phe Leu Leu Ala His Ala Phe Thr Glu  
465 470 475 480

Ala Leu Pro Gly Thr Pro Leu Gln Glu Arg Met Gly Glu Cys Leu Gln  
485 490 495

Arg Thr Gly Thr Ser Val Val Leu Thr Ser Ile Asn Asn Met Ala Ala  
500 505 510

Phe Leu Met Ala Ala Leu Val Pro Ile Pro Ala Leu Arg Ala Phe Ser  
515 520 525

Leu Gln Ala Ala Ile Val Val Gly Cys Thr Phe Val Ala Val Met Leu

530                    535                    540  
Val Phe Pro Ala Ile Leu Ser Leu Asp Leu Arg Arg Arg His Cys Gln  
545                    550                    555                    560  
Arg Leu Asp Val Leu Cys Cys Phe Ser Ser Pro Cys Ser Ala Gln Val  
565                    570                    575  
Ile Gln Ile Leu Pro Gln Glu Leu Gly Asp Gly Thr Val Pro Val Gly  
580                    585                    590  
Ile Ala His Leu Thr Ala Thr Val Gln Ala Phe Thr His Cys Glu Ala  
595                    600                    605  
Ser Ser Gln His Val Val Thr Ile Leu Pro Pro Gln Ala His Leu Val  
610                    615                    620  
Pro Pro Pro Ser Asp Pro Leu Gly Ser Glu Leu Phe Ser Pro Gly Gly  
625                    630                    635                    640  
Ser Thr Arg Asp Leu Leu Gly Gln Glu Glu Glu Thr Arg Gln Lys Ala  
645                    650                    655  
Ala Cys Lys Ser Leu Pro Cys Ala Arg Trp Asn Leu Ala His Phe Ala  
660                    665                    670  
Arg Tyr Gln Phe Ala Pro Leu Leu Gln Ser His Ala Lys Ala Ile  
675                    680                    685  
Val Leu Val Leu Phe Gly Ala Leu Leu Gly Leu Ser Leu Tyr Gly Ala  
690                    695                    700  
Thr Leu Val Gln Asp Gly Leu Ala Leu Thr Asp Val Val Pro Arg Gly  
705                    710                    715                    720  
Thr Lys Glu His Ala Phe Leu Ser Ala Gln Leu Arg Tyr Phe Ser Leu  
725                    730                    735  
Tyr Glu Val Ala Leu Val Thr Gln Gly Gly Phe Asp Tyr Ala His Ser  
740                    745                    750  
Gln Arg Ala Leu Phe Asp Leu His Gln Arg Phe Ser Ser Leu Lys Ala  
755                    760                    765  
Val Leu Pro Pro Pro Ala Thr Gln Ala Pro Arg Thr Trp Leu His Tyr  
770                    775                    780  
Tyr Arg Asn Trp Leu Gln Gly Ile Gln Ala Ala Phe Asp Gln Asp Trp  
785                    790                    795                    800  
Ala Ser Gly Arg Ile Thr Arg His Ser Tyr Arg Asn Gly Ser Glu Asp  
805                    810                    815  
Gly Ala Leu Ala Tyr Lys Leu Leu Ile Gln Thr Gly Asp Ala Gln Glu  
820                    825                    830  
Leu Leu Asp Phe Ser Gln Leu Thr Thr Arg Lys Leu Val Asp Arg Glu  
835                    840                    845  
Gly Leu Ile Pro Pro Glu Leu Phe Tyr Met Gly Leu Thr Val Trp Val  
850                    855                    860

Ser Ser Asp Pro Leu Gly Leu Ala Ala Ser Gln Ala Asn Phe Tyr Pro  
865 870 875 880

Pro Pro Pro Glu Trp Leu His Asp Lys Tyr Asp Thr Thr Gly Glu Asn  
885 890 895

Phe Arg Ile Pro Pro Ala Gln Pro Leu Glu Phe Ala Gln Phe Pro Phe  
900 905 910

Leu Leu Arg Gly Leu Gln Lys Thr Ala Asp Phe Val Glu Ala Ile Glu  
915 920 925

Gly Ala Arg Ala Ala Cys Ala Glu Ala Gly Gln Ala Gly Val His Ala  
930 935 940

Tyr Pro Ser Gly Ser Pro Phe Leu Phe Trp Glu Gln Tyr Leu Gly Leu  
945 950 955 960

Arg Arg Cys Phe Leu Leu Ala Val Cys Ile Leu Leu Val Cys Thr Phe  
965 970 975

Leu Val Cys Ala Leu Leu Leu Leu Asn Pro Trp Thr Ala Gly Leu Ile  
980 985 990

Val Leu Val Leu Ala Met Met Thr Val Glu Leu Phe Gly Ile Met Gly  
995 1000 1005

Phe Leu Gly Ile Lys Leu Ser Ala Ile Pro Val Val Ile Leu Val Ala  
1010 1015 1020

Ser Val Gly Ile Gly Val Glu Phe Thr Val His Val Ala Leu Gly Phe  
1025 1030 1035 1040

Leu Thr Thr Gln Gly Ser Arg Asn Leu Arg Ala Ala His Ala Leu Glu  
1045 1050 1055

His Thr Phe Ala Pro Val Thr Asp Gly Ala Ile Ser Thr Leu Leu Gly  
1060 1065 1070

Leu Leu Met Leu Ala Gly Ser His Phe Asp Phe Ile Val Arg Tyr Phe  
1075 1080 1085

Phe Ala Ala Leu Thr Val Leu Thr Leu Leu Gly Leu Leu His Gly Leu  
1090 1095 1100

Val Leu Leu Pro Val Leu Leu Ser Ile Leu Gly Pro Pro Pro Glu Val  
1105 1110 1115 1120

Ile Gln Met Tyr Lys Glu Ser Pro Glu Ile Leu Ser Pro Pro Ala Pro  
1125 1130 1135

Gln Gly Gly Gly Leu Arg Pro Glu Glu Ile  
1140 1145

<210> 2  
<211> 3453  
<212> DNA  
<213> HUMAN

<400> 2  
atccccggcca gcatgactcg atcggccccc ctcagagagc tgcccccag ttacacaccc 60  
ccagctgaa ccgcagcacc ccagatccca gctgggagcc tgaaggctcc actctggctt 120  
cgtgcttaact tccagggcct gctttctct ctggatgcg ggatccagag acattgtggc 180  
aaagtgtct ttctggact gttggcctt gggccctgg cattaggtct cgcacatggcc 240  
attattgaga caaacttggaa acagctctgg gtagaagtgg gcagccgggt gagccaggag 300  
ctgcattaca ccaaggagaa gctggggag gaggctgcat acacctctca gatgtgtata 360  
cagaccgac gccaggaggg agagaacatc ctcacaccccg aagcaacttg cctccaccc 420  
caggcagccc tcactgccag taaagtccaa gtatctactt atggaaatgc ctgggatttg 480  
aacaaaatct gctacaagtc aggagttccc cttattgaaa atgaaatgt tgagcggatg 540  
attgagaagc tggttccgtg cgtgatccctc acccccctcg actgcttctg ggagggagcc 600  
aaactccaag ggggctccgc ctacctgccc gcccggccgg atatccagtg gccaacctg 660  
gatccagagc agctgctgga ggagctgggt cccttgcct cccttgaggg cttccggag 720  
ctgctagaca aggcacaggt gggccaggcc tacgtggggc ggcctgtct gcaccctgtat 780  
gacctccact gcccacccat tgcccccaac catcacagca ggcaggctcc caatgtggct 840  
cacgagctga gtgggggctg ccatggcttc tcccacaaat tcatgactg gcaggaggaa 900  
ttgctgctgg gaggcatggc cagagacccc caaggagagc tgctgagggc agaggccctg 960  
cagagcacct tcttgctgtat gagtccccc cagctgtacg agcatttccg gggtactat 1020  
cagacacatg acattggctg gagtgaggag caggccagca cagtgtata agetctggcag 1080  
cggcgctttg tgcagctggc ccaggaggcc ctgcctgaga acgcttccca gcagatccat 1140  
gccttctccct ccaccacccct ghatgacatc ctgcacatgcgt tctctgaatg cagtgtgtcc 1200  
cgtgtgttgg gaggctatct gctcatgtg gcctatgcct gtgtgaccat gctgcgggtgg 1260  
gactgcgccc agtcccaggg ttccgtggc cttggccggg tactgtgtt ggcctggcg 1320  
gtggcctcag gccttggctg ctgtgcccgt ctcggcatca cttcaatgc tgccactacc 1380  
caggtgctgc ctttcttggc tctggaaatc ggcgtggatg acgtatttcc gctggcgtat 1440  
gccttcacag aggctctgcc tggcacccct ctccaggagc gcatggcga gtgtctgcag 1500  
cgcacgggca ccagtgtcgt actcacatcc atcaacaaca tggccgcctt cctcatggct 1560  
gccctcggtc ccatcccgtc getgcgagcc ttcccttccat agggggccat agtgggttggc 1620  
tgcacccctt tagccgtat gcttgccttc ccagccatcc tcagcctgga cctacggcgg 1680  
cgccactgccc agccgttgc tgcgtctgc tgcttctcca gtcctgtct tgctcagggt 1740  
attcagatcc tgccccagga gctgggggac gggacagtac cagtggcat tgcccaccc 1800  
actgcacacag ttcaagccct taccactgt aaagccagca gccagcatgt ggtcaccatc 1860  
ctgcctccccc aagccaccc ggtgccccca cttctgacc cactgggtcc tgagctttc 1920  
agccctggag ggtccacacg ggacccctta ggccaggagg aggagacaag gcagaaggca 1980  
gcctgcaagt ccctggccctg tgccctgtt aatcttgcctt atttcggccg ctatcagttt 2040  
gccccgttgc tgcttcagtc acatgctaag gccatgtgc tgggtctt tgggtctt 2100  
ctgggcttgc gccttacgg agccacccctt gtgcagacg gcctggccct gacggatgtg 2160  
gtgcctcggg gcaccaagga gcatgcctt cttgcgtccccc agctcaggta cttctccctg 2220  
tacgagggttgc ccctgggtac ccagggttggc ttgtactacg cccactccca acgcggccctc 2280  
tttgatctgc accagcgtt cttcccttc aaggccgtgc tgcccccacc ggccacccag 2340  
gcaccccgca cctggctgca ggcacccctt aactggctac agggaaatcca ggctgcctt 2400  
gaccaggact gggcttctgg gcatgccttcc cgcacactcg accgcacatgg ctctgaggat 2460  
ggggccctgg cctacaagct gctcatccag actggagacg cccaggagct tctggatttc 2520  
agccagctga ccacaaggaa tgcgttgc gctgggtgac agagaggagc tgattccacc cgagctttc 2580  
tacatggggc tgaccgtgtg ggtgagcagt gacccctgg gtctggcgc ctcacaggcc 2640  
aacttctacc cccccacccctt tgaatggctg caccacaaat acgcacccac gggggagaac 2700  
tttcgcatcc cgccagctca gccccttgc tttggccctgt tcccttcctt gtcgttggc 2760  
ctccagaaga ctgcagactt tgcgttgc tttggagggc atcgaggggg cccggccagc atgcgcagag 2820  
gccggccagg ctgggttgc tatctggggc tgcggcgctg cttccctgt ggcgttgc ttcgttgc 2880  
ctcgctgttgc ctctgtctgt cctcaacccctt tggacggctg gcctcatatgt gtcgttgc 3000  
gcatgtatgt cattgttgc gctgggttcc tttggatcc atgggttcc tgggtatcaa gtcgttgc 3060  
atccccgtgg tgatccctgtt ggcctctgtt ggcattggcg ttgagttcac agtccacgtg 3120  
gctctgggttcc tcctgaccac ccaggggcagc cggaaacctgc gggccggccca tggcccttgg 3180  
cacacatttgc cccccgttgc cgtatggggcc atctccacat tgcgtgggtct gtcatgttt 3240  
gctgggttccc actttgtactt cattgttgc gtttttttgc tggcgttgc agtgcgttgc 3300  
ctcctggggcc tcctccatgg actcgttgc tgcctgtgc tgcgttgc tgcgttgc 3360  
ccgcccaggagg tgatatacgat gtacaaggaa agcccaaggaa tcctgagtcc accagctcca 3420  
caggaggccggc ggcttagggcc cgaggagatc tag 3453

<210> 3  
<211> 126  
<212> DNA  
<213> HUMAN

<400> 3  
gagcgcattgg gcgagtgtct gcagcgcacg ggcaccaggta tcgtactcac atccatcaac 60  
aacatggcccg ctttcctcat ggctgccctc gttcccatcc ctgcgctgct agccttctcc 120  
ctacag 126

<210> 4  
<211> 119  
<212> DNA  
<213> HUMAN

<400> 4  
gcggccatag tggttggctg caccttgta gccgtgatgc ttgtcttccc agccatccctc 60  
agcctggacc tacggccggc ccactgcccag cgcccttgatg tgctctgtg ctctccag 119

<210> 5  
<211> 12886  
<212> DNA  
<213> HUMAN

<400> 5  
cggtgaatc ccggcgccgc gccccggacc cgcaactccc tgcactcctc cctcccgcc 60  
gcttaacac ccacacccca cagtctctcc cacgscggc cttggccgc cccactgaat 120  
ccctacgcgg ggcccagcgg taccgggaga cggggctagc ctatgggagc gcccagataa 180  
cgccgggttgg gggcgccgc gccccatcc cggccagcat gactcgatcg cggccctca 240  
gagagctgcc cccgagttac acaccccaag ctcgaaccgc agcaccagg gtgagtagag 300  
ggggagctgg aagaaggaag agagcggagc caggtctgtc actcgggct ctgcaagggtt 360  
tgtatgtct tgaagtgcgg agtgtcatta gatgtctgaa ggcaagttag agccagcacc 420  
gcaagcaagt tgtgcgtgtg tgtcggtgtg tctgtccgg tgcctccca tcgtctggc 480  
agtgagaatg aatgtctgtg gttcacccct tgcgtccacc cgacgacagg tgcgtgtaca 540  
tatgtatctt gcttcagaa aatgggccta tgccggccgg cgcgggtact cacgcctgt 600  
atcccaacac tggggaggctg aggcaaggcattt attacctgat gtcaggagtt cgagaccgc 660  
caggccaaaca tggggaaact ctgtctctac taaaaataaa aattagcagg gctgggtggc 720  
ggcgccctgt agtccaact actcgggagg ctgaggcagg agaatctttt gaaacctggg 780  
ggcggagggtt gcagtcaga cggatcaca ccactgcact ccagccagg caacagagcg 840  
agatgcgtct caaaaaaaaaaaa aaaaaaaaaaaa aaaaggagag aaaacaaaaaa gaaaagaaaag 900  
gaaaataggc ctatgccttc ctcaggtgtg tgctggggat ggtgggtgtt acatcttcca 960  
agtctggggc tgtgtctgtg ttgggtctcc ctgtccaca tccagaaatc aagaagcgag 1020  
ggctgggcag cagatataca gggtgagaag ggaaggattt catgcattgt tacagtgtat 1080  
cctggctgac ctttcctttt ccatcccaaa tccatgtgg gacgcctgaag gctccactct 1140  
ggcttcgtgc ttacttccag ggcctgtct tctctctggg atgcgggatc cagagacatt 1200  
gtggcaaagt gctttctg ggactgttg ctttggggc cttggcatta ggtctccgca 1260  
tggccattat tgagacaaac ttggAACAGC tctgggtaga aggttaagtgg tggacactgg 1320  
ccatagctgc tcaggtatgg tgagcccaag acaagaacgg ggtgaggagc tgctactga 1380  
gctctagcag gcctggccct ggggctggag ggtcacctgc tggccagagt cctacacctg 1440  
ggcatctagg aacacggttc cttccaccaa ccataaatgg acatctgcag atgttagacc 1500  
ttgtgaattt ggtgtataga atattgtct gtgattttgc ttaagcccta gaccacacaa 1560  
gaaatccgca gagccacaca accgaggctt ctccacttcc ctggcagagc tccaggaccc 1620  
agaagcctgg ggagtcggcc tggggccatga ccaggccctc agggataagg cagggactcc 1680  
tctctggcc accttccttca cttgtctgg gaggcctcag agttagaagg ccccttcgca 1740  
agttctgacc cgccttagggc ctttcccccc cagtttctc tggctgtct ggcagtggc 1800  
agccgggtga gcccaggatc gcattacacc aaggagaagg tgggggagga ggctgcatac 1860  
acctctcaga tgctgatata gaccgcacgc cagggggag agaacatctt cacacccgaa 1920  
gcacttggcc tccacctcca ggcagccctc actgcccagta aagtccaaatc atactctat 1980

ggaagggtga gtctggctga gccccctgagc agctgggggc gagggcgtgct gtgggggttc 2040  
tggagtggga atccccttct tctgctgatc tccttatgccctt ctggctattt cagtcctggg 2100  
atttgaacaa aatctgctac aagtcaggag ttccccttat taaaaatggaa atgattgagc 2160  
gggttaagtgt cctgagaggg agtagaggca gaacttttc tgtagcgtgg gaggactcag 2220  
agaccgagca agcccccacag cctgcaatct gcccccttaa aactaaggag ggggattgca 2280  
gagggcatcc tacaaagggtt gtggggcagg actgacgtgg cccggggtat ccctggcaga 2340  
tgattgagaa gctgttccg tgcgtgatcc tcaccccccct cgactgcttc tgggagggag 2400  
ccaaactcca agggggctcc gcctacctgc cgtgagtgcc actcctgggg ccctgttca 2460  
tctcccgctg gggactctcc cagcagaaaag gagggtctg ggaatgagg atgatcaaaa 2520  
ccttaccaag gtcttaatta cctcccaaggc caggaacaga gacatgggc ttccccaaagg 2580  
ctctctccac atctcccttc tctttccctc tcaaggaagg aagacctgac ttatttacac 2640  
aaaactaaac acaaagatct gtaagatctg agcaaaggag aaaaagatcc ccacaaagag 2700  
gcttgcgtgg gggaaattac ctaggtgtt gctaagccat tgcccaaggcc agaaagaaaa 2760  
cctgctacag gcatgtgcct gctgggtgtt tattagaacc aagcacacag ctggtaagg 2820  
aactcagtgg ggcccttctg ggcccttctt atgttattagg taaccctgccc ctgatattcg 2880  
tctcagcccc ttgtactctt ctacagctca ctgttagcacc ctgggtggcc catgcagcct 2940  
ggcagttctg agaagctgag gcttgacac cctccatatg gaaggacaaa tcggcagata 3000  
agaggagggt ggggtacagc atggcgcfffff agcagcagg tggagcctgg gtttgcgtcc 3060  
ctgaccctca ccaactatag gctttccct cagcggccgc ccggatatcc agtggaccaa 3120  
cctggatcca gagcagctgc tggaggagct gggtccctt qcctcccttq agggcttccg 3180  
ggagctgcta gacaaggcac aggtggccca gcctacgtg gggccggccct gtctgcaccc 3240  
tcatgacccctc cactgccccac ctatgtcccc caaccatcac agcagggcagg tgggttccaa 3300  
ccaggtctgc cagggaaagg ctgtttccct tccctttccc ttccctcatac tcctgtgttc 3360  
tgggggagct gactgctctg tgccctgacc ccccaacttcc tggccattat taccctgtctc 3420  
ccacagtgcc agggcccaa tgffccatttcc cattcaggat atccctacggg gccctcaagt 3480  
ggtatatatg aatccctttt tcctttcttta agccttagata aggctggact tctttttttt 3540  
ttttttttt agtctcaactc tgcacccctt gctggaggtc agtagttcg tcttggctca 3600  
ctgcaacctc ggctcaagca attctcttgc cttagcttcc ttagtagctg ggattacagg 3660  
tgccccaccac catgccccggc taattttat tagccttccca aagtgttggg attacaggcg 3720  
tgagccactg cgcctggccca aggctggact tttatcaaa atagactaat acaggggaaac 3780  
taagaacaca gcaggttaagc atgaatatca tacctgggtt cccaggtttc tttgtggccc 3840  
tgcaaatgtg gtactttttt cagaatccgc cagtttacacc agctctccc agaaggctac 3900  
ttccagggct ctgtttcccc ttggggcttcttgcgtt gatactagct gttcactctt 3960  
gcagagcagt caagaggctc agaatagttt cctactacc agccctactg agcttcatgg 4020  
cagcgtgggtt cctggaggtt gaagccagg gacactcaat tatccacggc cagggccctt 4080  
agcattaaacc cctctgttcc ccctccaggc ctcccaatgt ggctcacgg ctgagtgggg 4140  
gctgccccatgg ctctcccaaaattctatgc actggcaggaa ggaatttgcgtt ctgggagggca 4200  
tggccagaga ccccaagggaa gagctgtga gttttttttt cctctggaggg 4260  
gactctgttcc atgagaaccc atactgttaat gccaggcagc tctggcaaaa ggccttcac 4320  
atcccttacc agtgttttttgg gccagctctg accccctgggtt ctcccacacc cccaccagg 4380  
cagggggctt gcaagagcacc ttcttgcgtt gtagtcccccc ccagctgtac gaggatttcc 4440  
gggggtgacta tcagacacat gacattggctt gtagtggggaa gcaggccagc acagtgttcc 4500  
aaggcctggca gggcgctttt gtgcaggctcg gtatggacaa ggacaagggg ggtgccttga 4560  
ggccatcccc tcttctgtcc ccctcttatac caccctgtt ctccagctgg cccaggaggc 4620  
cctgcctgag aacgcttcccc agcagatcca tgccttcttcc tccaccaccc tggatgacat 4680  
cctgcatgcg ttctctgttgc tcaatgtgtc ccgtgtgggtt ggaggctatc tgctcatgg 4740  
gggtctgtca cttggcacct tggcccccacc ccaccccttccaa ccagtgccca ccctgggagc 4800  
ccctgtgagact gccccttccc cccacagctg gcctatgtctt gtgtgaccat gtcgggtgg 4860  
gactgcccccc agtcccagggtt tccagggtgtt gccccttcttgc gctctggaa ggcctggcg 4920  
gtggcccttag gctttgggtt ctgtggcccttgc tccggcatca ccttcaatgc tgccactacc 4980  
caggtacgccc agactgca ggcagactca gtggcaggatcc ccaggcttca cgggtcttca 5040  
gctgccccctt cctctgtcccc tccagggtgtt gccccttcttgc gctctggaa tggcgttgg 5100  
tgacgtattt ctgtgtggcc atgccttacca agaggctctg cctggcaccc ctctccagg 5160  
ggggccttgc tcccccaaggcc tcatctgagg cagctcagct tactggttaa gagccttctt 5220  
gttcaagtga cccttgggtt gctaatgttac ctcgggtgtt cttgtccccca tctgtaaaca 5280  
ggggaaataa tagtgtgtgtt tccttaagggtt tattgtttgg atcagtgttgg taactcaagt 5340  
tgaatgttta gaaacagcccc tcatacgtac atggtaccca ataaatgttta gccactgtgt 5400  
tatgactgcc ccacctctgc accccaaagtt cctgagccctc cccttcactc cactttgaca 5460  
cggcccttcc cttgtgaccc tggggcaggat ccccaactctg tcctggcagg agcgtatggg 5520  
cgagtgtctg cagcgcacgg gcaccaggat cgtactcaca tccatcaaca acatggccgc 5580  
cttccttcatg gtcgttcccttgc ttcccatcccc tgcgttgcga gccttcttcc tacaggccgc 5640

catatgggtt ggctgcacct ttgttagccgt gatgcttgtc ttecccagcca tcctcagcct 5700  
ggacctacgg cggcccaact gccagcgctc tgatgtgctc tgctgtttct ccaggtactg 5760  
cgtgcgcccc agcccccttc tcccgtgacc cacggccagcc tgccccctca ccagcatttc 5820  
aaggcacaga cctgtcatcc actctctacc tcttccagtc cctgctctgc tcaggtgatt 5880  
cagatcctgc cccagagact gggggacggg acagtaccag tgggcatttc ccacccact 5940  
gccacagtcc aagccttac ccactgtgaa gccagcagcc agcatgtggt caccatcctg 6000  
cctccccaag cccacccgtt gccccccact tctgaccac tgggctctga gctttcagc 6060  
cctggagggt ccacacggga cttcttaggc caggaggagg agacaaggca gaaggcagcc 6120  
tgcaagtccc tgccctgtgc ccgctggaat cttgcccatt tcgcccgtca tcagtttgcc 6180  
ccgttgcgtc tccagtcaca tgctaaggta agactggca gagcaggggca gagacttagc 6240  
atctctggc ccagaaggcc agagagggt tagtccactg cctgaggggc tggggcagc 6300  
cctggggctc ccagcttagt tgctcatacc cgcaggccat cgtgctgggt ctctttgggt 6360  
ctcttctggg cctgagcctc tacggagcca ctttggcga agacggccctg gcctgtacgg 6420  
atgtgggtcc tcggggcacc aaggagcatg ctttcttgag cggccagtc aggtacttct 6480  
ccctgtacga ggtggccctg gtgaccacgg gtggcttga ctacgcccac tcccaacgcg 6540  
ccctcttta tctgacccagg cgcttcaggta ccctcaaggc ggtgctggcc ccacccgcca 6600  
cccaggcacc cccgacccctgg ctgcactatt accgcaactg gctacagggt gagaggcgag 6660  
gagacgggca gggaggggtg ctgcaggggag aaacgcctg gggccacccag ctaatagaac 6720  
cctatctgg tctccccccag gaatccagc tgcccttgac caggactggg ctcttggcg 6780  
catcacccgc cactcgaccg caatggctt gaggatgggg ccctggccta caagctgctc 6840  
atccagactg gagacgcccga ggagcttctg gattcagcc aggttggag agggtggag 6900  
gggtccacta gtacaggggc tgcaggccctc ctggggccag gccttcagcc ctctctgcct 6960  
ctgcagctga ccacaaggaa gctgggtggac agagaggac tgattccacc cgagctcttc 7020  
tacatggggc tgaccgtgtg ggtgagcgt gacccttgg gtctggcagc ctacaggcc 7080  
aacttctacc ccccacctcc tgaatggctg cagcacaat acgacaccac gggggagaac 7140  
tttcgcagtg agtcttgggg ggagctcgcc aagagcctca gcctcgccca cacaagccct 7200  
gagcctgagg ccctgcccac tctccccctg gtcaccgccc ctgtccctct ccctttctc 7260  
ccttccccctc ccctccacag tcccggcagc tcagccctt gagtttggcc agttcccttt 7320  
cctgctgcgt ggcctccaga agactgcaga ctttggag gccatcgagg gggccggggc 7380  
agcatgcga gaggccggcc aggctggggt gcaacgcctac cccagcgct ccccttctc 7440  
cttctggaaa cagtatctgg gcctgcggcg ctgttctgt ctggccgtct gcatcctgtc 7500  
gtgtgcact ttccctgtct gtgtctgtc gtcctcaac ccctggacgg ctggcctcat 7560  
agttagtgc tgccaggatg gggacagaga caccggccaccc ttccctggcc agctgtcat 7620  
ccctcctgcc aggacccctc tttggccct gtctccctca ggtgctgtc ctggcgatga 7680  
tgacagtgga actcttttgt atcatgggtt tcctgggcat caagctgagt gccatcccc 7740  
tggtgatctc tttggccctc gttaggcattt gctttttttt gacgtccac gtggctctgg 7800  
tgagcacggg caccggggg agggaccaat cagctgattt agtattcaac acatattttt 7860  
caagccccca ctatgtgta ggtactattt aagaattttt ggtgggttga cgtggtagct 7920  
cattcctgtt atccccagcac tttggggaggc cgaggcaggat ggatcacctg aggtcaggag 7980  
ttcgaacacca gcctggccaa catggtgaaa ccctgtcttt actaaaaata caaaaaatta 8040  
gccaggcggtg ttggccatcg ccagtaatcc cagctacttt ggaggctgag gcagaatttc 8100  
ttgaacctgg gaggcgaagg ttgcagtgtt ctgagatcg gtcattgcac tccagcctgg 8160  
gcaacaagag taaaacttc cgtctaaaa aaaaaaaaaaaa aaaagaattt gggctggca 8220  
cagtggctca tgcctgtat tggatgtat ctggggcatt ttgggaggcc aaggcaggcg 8280  
gatccccctga agttagggat tcaagaccag cttggccaaac attgcaaaac cccgtctcta 8340  
ctgaaaatac aaaaattttt tgggcgtgtt ggctcatgcc tgtaatccca gctactcagg 8400  
aggctgaggc aggagaattt cttgaaccca ggaggcggag gttacagtga gctgagatca 8460  
catcaactgtt ctccagccctg ggcaaaagag caagattcaa tcttcaaaaa aagaattttt 8520  
aaaataaaaaaaa taaaagaat acgggatata atagcaatac agttttttac ctccaaggaa 8580  
cttatattct agggatagag atagacaata agggctgggt gaggtggctt acgcctgtaa 8640  
tcccagcact ttgggaagcc gaggtgggca catcaactgtt ggtcaggagt tcaagaccag 8700  
cctggccgac atggtaaac cccatctca cttttttat gaaaattttt tgggtgtggg 8760  
ggtgcatgcc tgtaatccca gctacttggg aggtgaggcc aggagaatca cttgaaccccg 8820  
ggagggtgga ggttgcctt agccgagacc atgccactgc actccagcct gggagaccga 8880  
gcaaaaactcc atctcaaaaa aaaaaaaaaaaa aacctcagcc tcccaaaatgt ctgggattac 8940  
aggcatgagc caccgtggcc gggcttggg ttttaagag acagggtctc actctgtcac 9000  
ccaggctgtt cccttatttc tacgtgtctc tttgtcttaag ctcaccaaaac ccatccatac 9060  
agtgttaccc aaaaggactc acgcatgctc cctgtacagt tcccatgaac ttatgttacc 9120  
cagccttgg tggccaagg tatgttggag ctctgaagac agatacagct cgggacatgg 9180  
ctgagctggc catgactggc agaggagcag ctccaggacc actctttt ccttagggctt 9240  
cctgaccacc caggcagcc ggaacctgctg gggcccat gcccatttgc acacatttgc 9300

ccccgtgacc gatggggcca tctccacatt gctgggtctg ctcatgcttg ctggttccca 9360  
ctttgacttc attgttaaggt agggagggct cggggcaggg aggcaggcgt caggacaggc 9420  
ctgggctgac tccccccaca cccttacccct aggtacttct ttgcggcgct gacagtgtc 9480  
acgctcctgg gcctcctcca tggactcgtg ctgcgtcctg tgctgctgtc catcctggc 9540  
ccgcccggcag aggtgaccac accctcgga ccatccctct actcccaagcc caagggacgg 9600  
gttagggaga ggcaaggaa gggacagagc cctgtggccc acagacaggt acctcccaa 9660  
caggtgccac cagctgaagg tggcagccctc ctcccttccc cagacaccat gttcctgccc 9720  
ctcagccctc ctggcttctt catgggaccc accttagact tttaggatcc agaacaagg 9780  
gcaggggttg ccccaggcct caacatccctg tcgcctgcca gctctcatat cctgctggag 9840  
accaacaagg gccccagctt cccaacagtc atggtaatcc ccagcgagat gctaaagggg 9900  
acgggagccc cagggggcccg tggcttact ggggctggtg tctcccaaca ggtgatacag 9960  
atgtacaagg aaagcccaga gatcctgagt ccaccagtc cacaggagg cgggcttagg 10020  
tggggggcat cctcccccct gccccagagc tttgcccagag tgactaccc tcgtaccgt 10080  
gccatccacc caccggggctt gcctggtgcc tacatccatc cagccctgtc tgagccccc 10140  
tggtccccctg ctgtcaactag ctctggcaac ctcagttcca ggggaccagg tccagccact 10200  
gggtgaaaga gcagctgaag cacagagacc atgtgtgggg cgtgtgggtt cactggaaag 10260  
caactgggtct ggtttagac gcaggatgga cccctggagg gctctgtc tgctgatcc 10320  
cctctcccgaa cccagctgtc atgggcctcc ctgtatccca tacagaacag ccaccgattt 10380  
gcacatccag gcctgtgtga gcctgtatet gtgtcaactg agagtaaaag ctggcaactg 10440  
gggctgcagt gcagccctgt ccccttccc accccacacc actgcctgcc cagctgacca 10500  
agcctgaggg accctccagc accctccctg ctggtgactc ctggcaggc tctccatata 10560  
cctggccacc tcctaccaca tccatttattt atatgaaaat gtctattttt gtgtatata 10620  
tacatgttag ctatgtgaa agttttattt tttaaagaat gaaatatattt gaaatatattt 10680  
ctatgtgaa agttttattt tttaaagaat gaaatatattt ctatgtgtc 10740  
agttttattt tttaaagaat gaaatatattt ctatgtgtc aagtgaacat tagcttcagt 10800  
tgctttttt tggacagagt ggggagtttgc caagtgaaca ttagctattt gaaggagctt 10860  
ctctgggtcc aggacctgag gtattagctt ctctagttct ggggtggaaa gaccccgat 10920  
tctggatttt tgcataatac ttggtaacat cacttggatt aagtcttac tataaaaaac 10980  
gataacaat ttgttgggtg taaaatctt ctgggttcaa tctggagacc gagagcagaa 11040  
aaaaaagaac cccactgtgt ggcttcaga gccaccatata tccagctgc ccgtctctcc 11100  
agactcacct ccacccatct gcttcacccg cacgggaaac ggcaaggcag aggggcaaaag 11160  
ccatgcagca ggtgaaggc gaggtggagg cagatcagga aagcagccag ttgaagcaga 11220  
gagaggctaa caggtctgg ggagcttctc aggagggtttg tggaccagg gaaaggagcc 11280  
aggttccaga gcaaccttca aggcaaaggc ctctgttaagt tggttgtcct gacagccgag 11340  
aggtgtctt ggcctgtcgt ccagtggatc agttgcggga actgctcaga aactgagggt 11400  
ctagcagttt gtgaggacac agcgttaagtt gtttggatc tggaaaggta acagctccac 11460  
taagcagagg ctttgcagag tggccacacgc ctttgcataatc agcacagagc ctcacccat 11520  
ggcgtgggaa ggttgcac tggcccttcc cagccatagc ttaggaccca tagtcttagt 11580  
cacatagacc ctggcttcca accacccact caccaggaat gatcccaccc caggaacaaat 11640  
gcgttctcac atcccacccccc acctggacaa agggcaggaa atcatgttct gacccaaaga 11700  
tacaacaaca aaaacaacaa caacaaaaaa cgcttattgc aattgaatcc acgctaaaat 11760  
gcctaaaaag ctcaagagaa gcgggttagtt ggcagagaac ctagagtagg ggtgtcaacc 11820  
agcaggccca agggagggag gctgcattt ggttccagcag tggttggtc accaagaagg 11880  
gccttcttagg tggagcagag agagctcacc agggcagaat agtgc当地 ggtcagccc 11940  
tcagtgccac ttaccagcgg agtaaccctg ggcaagtttgc ccagccctac taagcctccc 12000  
cattttcattc ttccaggccc cgaggagatc tagcctctgc ctcccccccc agcaccctt 12060  
catcagacac aaggagcgtt actgtctggc caggctgaat tggctctcgg gttccctaaat 12120  
tctcatacgc cattttcttgc ctcttgcataacc ctttgcattc tccccctgtat gtgaccccat 12180  
atcacccttc gaggtgaatt ggatcggatc ccatcttctc caggagggtt ggggtcgtgc 12240  
ctccctgttagt gtccctgtcgt ccctgttagt ctgtggccgt ctgtttcccc gttccctctct 12300  
ctaagccccc aggcttactg cgggtaaaggc cggccggaca ggacctaacc ccctgggacg 12360  
aacaccagct ccgcaaaggc ctccgcaccc ggcccccggcc acgggggtgcg ggtcccagga 12420  
ggaccagcag agaggagcat aggagagcaa aggagatcag tgaccatgg ctccccccgg 12480  
ggcgtggaaac agcccgagc cgcctgtgtat ttgcataacc atgggtgcacc acgaaaaagat 12540  
acccctcaaga tgcctgtcact ccctctgtgc gcgcatttct gcactgtttt agagcatgat 12600  
gccttcttaca cgcattgtg tgcataaaact acatataagg agtgcgtacc acgcaggcat 12660  
ccaacaacca taagtgttgtt aagtgttagt tctccctgca aggttgcag cggaaaggtcac 12720  
gaatataactc gggtttctct tcaaagcgtca taaatcttgc gcttttact aaagatttcc 12780  
gtggagagaa agttgtgagt ttttatttcaaa tttttgagg cctcttattt cctgaggctt 12840  
cattttaag tattaaagg taggcaacta caaaaaaaaaaaaaa 12886

## INTERNATIONAL SEARCH REPORT

1

International application No.

PCT/SE 99/01784

## A. CLASSIFICATION OF SUBJECT MATTER

**IPC7: A61K 48/00, C07K 14/705, A61K 38/17, G01N 33/68**  
 According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

**IPC7: A61K, C07K, G01N**

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

**SE,DK,FI,NO classes as above**

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 9929854 A1 (ONTogenY, INC.), 17 June 1999 (17.06.99), page 73 - page 77, claim 3 --	1-12,16-22
P,X	EP 0879888 A2 (SMITHKLINE BEECHAM PLC), 25 November 1998 (25.11.98), claim 9b --	1-12,16-22
P,X	Journal of Clinical Investigation Online, Volume 95, No 23, November 1998, David Carpenter et al., "Characterization of two patched receptors for the vertebrate hedgehog protein family", pages 13630-13634 --	1-12,16-22

 Further documents are listed in the continuation of Box C. See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

- "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

- "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

- "&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

1 February 2000

14-02-2000

Name and mailing address of the ISA/  
 Swedish Patent Office  
 Box 5055, S-102 42 STOCKHOLM  
 Facsimile No. + 46 8 666 02 86

Authorized officer

Patrick Andersson/AE  
 Telephone No. + 46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 99/01784

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Nature Genetics, Volume 18, February 1998, Motoyama J. et al., "Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog", page 104 - page 106, figure 1, EMBL Acc no AB000847  --	1-12,16-22
A	FEBS Letters, Volume 410, 1997, Takashi Takabatake et al., "Hedgehog and patched gene expression in adult ocular tissues" page 485 - page 489  --	1-12,16-22
A	WO 9745541 A2 (THE LELAND S. STANFORD JUNIOR UNIVERSITY), 4 December 1997 (04.12.97)  -----	1-12,16-22

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/SE99/01784

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: Claims 12 partially and claim 21 completely because they relate to subject matter not required to be searched by this Authority, namely:

See next sheet.

2.  Claims Nos.: 13–15 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

See next sheet.

3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**  

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/SE99/01784

Box I. 1.

Claims 12 partially 21 completely relate to methods of treatment of the human or animal body by therapy/ diagnostic methods practised on the human or animal body See PCT Rule. 39.1.(iv). Nevertheless, a search has been executed for these claims. The search has been based on the alleged effects of the compounds.

Box I. 2.

Claim 13 relates to a screening method, it is not clear from this claim or the description how the polypeptide is used to achieve the screening. Claim 14 relates to a method of synthesis a modified drug, it is not clear from the description or the claim how the protein is used. Claim 15 relates to a drug identified by any of the methods in 13 or 14. Thus, claims 13-15 is considered to fail to comply with PCT-article 6 and have not been searched. Moreover, claim 15 concerns a drug per se, the drug may be a known substance. The description or the claims does not give any indication on that any unknown substance(s) that are preferred. For this reason, a search or a statement concerning novelty and an inventive step can not be made for claim 15.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE99/01784

Claims 12 partially 21 completely relate to methods of treatment of the human or animal body by therapy/ diagnostic methods practised on the human or animal body See PCT Rule. 39.1.(iv). Nevertheless, a search has been executed for these claims. The search has been based on the alleged effects of the compounds.

Claim 13 relates to a screening method, it is not clear from this claim or the description how the polypeptide is used to achieve the screening. Claim 14 relates to a method of synthesis a modified drug, it is not clear from the description or the claim how the protein is used. Claim 15 relates to a drug identified by any of the methods in 13 or 14. Thus, claims 13-15 is considered to fail to comply with PCT-article 6 and have not been searched. Moreover, claim 15 concerns a drug per se, the drug may be a known substance. The description or the claims does not give any indication on that any unknown substance(s) that are preferred. For this reason, a search or a statement concerning novelty and an inventive step can not be made for claim 15.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

02/12/99

International application No.

PCT/SE 99/01784

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9929854 A1	17/06/99	AU 1716599 A	28/06/99
EP 0879888 A2	25/11/98	CA 2232808 A GB 9710752 D JP 11075874 A GB 9805954 D	23/11/98 00/00/00 23/03/99 00/00/00
WO 9745541 A2	04/12/97	AU 3227497 A	05/01/98